

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2003, 10:41:09 ; Search time 38 Seconds
(without alignments)
2262.514 Million cell updates/sec

Title: US-10-043-787-1

Perfect score: 2271

Sequence: 1 MSDKLPYKVDIGLAAGRK.....QAQYLGMSCDGPFKPDHYRY 432

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 724715 seqs, 199017464 residues

Total number of hits satisfying chosen parameters: 724715

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2266	99.8	457	9	US-09-925-301-1059
2	1748	77.0	437	12	Sequence 1059, Ap
3	1733	76.3	433	12	Sequence 5097, Ap
4	1733	76.3	433	12	Sequence 2435, Ap
5	1633	71.9	449	12	Sequence 2430, Ap
6	1594	70.2	447	12	Sequence 1740, Ap
7	1546	68.1	426	12	Sequence 3683, Ap
8	1514.5	66.7	516	12	US-10-369-493-10135
9	1465	64.5	428	12	Sequence 13111, A
10	1428	62.9	468	12	Sequence 10445, A
11	1422	62.6	468	12	Sequence 9241, Ap
12	1406	61.9	476	12	Sequence 9386, Ap
13	1405	61.9	466	12	Sequence 8266, Ap
14	1401	61.7	469	12	Sequence 8750, Ap
15	1401	61.7	469	12	Sequence 15745, A
					Sequence 16129, A

16	1399	61.6	463	12	US-10-369-493-19704
17	1393	61.3	469	12	US-10-314-657-71
18	1372	60.4	485	15	US-10-156-761-12587
19	1357	59.8	461	12	US-10-369-493-21032
20	1355	59.7	461	12	US-10-369-493-4783
21	1355	59.7	461	12	US-10-369-493-7543
22	1342	59.1	446	12	US-10-369-493-17552
23	1340.5	59.0	463	12	US-10-369-493-7922
24	1336	58.8	459	12	US-10-369-493-10671
25	1311	57.7	446	12	US-10-369-493-15377
26	1310	57.7	463	12	US-10-369-493-20829
27	1302.5	57.4	466	12	US-10-369-493-19331
28	1300	57.2	465	12	US-10-369-493-17934
29	1296	57.1	461	12	US-10-369-493-11466
30	1296	57.1	461	12	US-10-369-493-14652
31	1296	57.1	464	12	US-10-369-493-15125
32	1296	57.1	466	12	US-10-369-493-12110
33	1256.5	55.3	463	12	US-10-369-493-16776
34	1242	54.7	478	11	US-09-746-660A-98
35	1242	54.7	498	11	US-09-919-854-2
36	1238	54.5	474	10	US-09-738-626-4336
37	1225.5	54.0	469	12	US-10-369-493-13881
38	1160	51.1	416	12	US-10-369-493-14175
39	1151	50.7	614	9	US-09-782-051-2
40	1098	48.3	432	11	US-09-746-660A-102
41	943	41.5	418	12	US-10-369-493-90
42	927.5	40.8	399	12	US-10-369-493-22854
43	885	39.0	416	12	US-10-369-493-9005
44	878.5	38.7	422	12	US-10-369-493-21670
45	871.5	38.4	425	12	US-10-369-493-1245

ALIGNMENTS

RESULT 1

US-09-925-301-1059
; Sequence 1059, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1059
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1059

Query Match 99.8%; Score 2266; DB 9; Length 457;
Best Local Similarity 99.8%; Pred. No. 3.6e-220;
Matches 431; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSDKLPYKVDIGLAAGRKALDIAENEMPGLMRERYASAKPLKGARIAGCLHMTVET	60
Db	26	MSDKLPYKVDIGLAAGRKALDIAENEMPGLMRERYASAKPLKGARIAGCLHMTVET	85
QY	61	AVLIETVLTGAEVQWSSCNIFSTQDHAATAAKAGIPYAWKGETDESYLWCIEQTLVF	120
Db	86	AVLIETVLTGAEVQWSSCNIFSTQDHAATAAKAGIPYAWKGETDESYLWCIEQTLVF	145
QY	121	KDGLNMLDDGGDLTLNHTKYPQLLPGIRISBETTTGVNLYKMANGILKVPAINV	180
Db	146	KDGLNMLDDGGDLTLNHTKYPQLLPGIRISBETTTGVNLYKMANGILKVPAINV	205

QY 181 NDSVTSKFNLYGCRSLIDIGIKRATDVNMAGKAVVAGYGVGKCAQALRGFGARVI 240
DB 206 NDSVTSKFNLYGCRSLIDIGIKRATDVNMAGKAVVAGYGVGKCAQALRGFGARVI 265
QY 241 ITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIG 300
DB 266 ITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIG 325
QY 301 HFVEIDVKWLNENAVEKNIKQVDYRLKNGRRILLAEGRVLNLCAMGHPSPVMSN 360
DB 326 HFVEIDVKWLNENAVEKNIKQVDYRLKNGRRILLAEGRVLNLCAMGHPSPVMSN 385
QY 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKKLDEAVAHAHLGKLNVLTKLTEKQAQYLGMS 420
DB 386 SFTNQVMAQIELWTHDPKYPVGVHFLPKKLDEAVAHAHLGKLNVLTKLTEKQAQYLGMS 445
QY 421 CDGPFKPDHRY 432
DB 446 CDGPFKPDHRY 457
RESULT 2
US-10-369-493-5097
; Sequence 5097, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5097
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5097

Query Match 77.0%; Score 1748; DB 12; Length 437;
Best Local Similarity 76.7%; Pred. No. 8.2e-168;
Matches 332; Conservative 41; Mismatches 56; Indels 4; Gaps 2;
QY 4 KLPYKVDIGLAAGRWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVETAVL 63
DB 5 KPAYKVDIKLADFGKKEIILAENEMPGLMRMRYSASKPLKGARIAGCLHMTVETAVL 64
QY 64 IETLVLTGAEVQWSSCNIFSTQNHAAAAIAKAGIPYVANKGETDEBYLWCIEQTLVFKDG 123
DB 65 IETLVLTGAEVQWSSCNIFSTQDHAAAAIAQTGPVYVANKGETDEBYLWCIEQTLVFKDG 124
QY 124 -PLANMILDGGDLTNLIHTKYPOLLPGIRGISETTTGVHNLKYMANGILKVPAINVND 182
DB 125 QPLNMILDGGDLTNLIHTKYPOLLPGIRGISETTTGVHNLKYMANGILKVPAINVND 184
QY 183 SVTKSFDNLYGCRSLIDIGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVIIT 242
DB 185 SVTKSFDNLYGCRSLIDIGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVIIT 244
QY 243 EIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHF 302
DB 245 EIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHF 304
QY 303 DVEIDVKWLNENAVEKNIKQVDYRLKNGRRILLAEGRVLNLCAMGHPSPVMSNSF 362
DB 305 DCEIDVKWLNENAVEKNIKQVDYRLKNGRRILLAEGRVLNLCAMGHPSPVMSNSF 364

QY 363 TNOVMAQIELWTH---PKYVPVGVHFLPKKLDEAVAHAHLGKLNVLTKLTEKQAQYLG 419
DB 365 TNOVMAQIELWTKFGTPQBYKGLVYLPKTLDEEVAYLHLAGVKLTLSDEQASYLG 424
QY 420 SCDGPFKPDHRY 432
DB 425 PVAGPYKPDHRY 437
RESULT 3
US-10-369-493-2425
; Sequence 2425, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2425
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2425

Query Match 76.3%; Score 1733; DB 12; Length 433;
Best Local Similarity 77.1%; Pred. No. 2.7e-166;
Matches 330; Conservative 40; Mismatches 56; Indels 2; Gaps 2;
QY 7 YKVADIGLAAGRWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVETAVLIET 66
DB 6 YKVADISLAAGRWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVETAVLIET 65
QY 67 LVTGAEVQWSSCNIFSTQNHAAAAIAKAGIPYVANKGETDEBYLWCIEQTLVFKDG-P 124
DB 66 LVALGAEVQWSSCNIFSTQDHAAAAIAATGVVPFAWKGETEYELWCIEQTLVFKDG-P 125
QY 125 LNMILDGGDLTNLIHTKYPOLLPGIRGISETTTGVHNLKYMANGILKVPAINVND 184
DB 126 LNMILDGGDLTNLIHTKYPOLLPGIRGISETTTGVHNLKYMANGILKVPAINVND 185
QY 185 TSKSFDNLYGCRSLIDIGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVIIT 244
DB 186 TSKSFDNLYGCRSLIDIGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVIIT 245
QY 245 DPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHF 304
DB 246 DPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHF 305
QY 305 EIDVKWLNENAVEKNIKQVDYRLKNGRRILLAEGRVLNLCAMGHPSPVMSNSFTN 364
DB 306 EIDVAWLKANAKADVNIIPQVDYRLKNGRRILLAEGRVLNLCAMGHPSPVMSNSFTN 365
QY 365 QVMAQIELWTHDPKYPVGVHFLPKKLDEAVAHAHLGKLNVLTKLTEKQAQYLGMS 424
DB 366 QVMAQIELWTHDPKYPVGVHFLPKKLDEAVAHAHLGKLNVLTKLTEKQAQYLGMS 425
QY 425 FKPDHRY 432
DB 426 YKADHRY 433
RESULT 4
US-10-369-493-2430
; Sequence 2430, Application US/10369493

Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 2430
LENGTH: 433
TYPE: PRT
ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2430

Query Match 76.3%; Score 1733; DB 12; Length 433;
Best Local Similarity 77.1%; Pred. No. 2.7e-166; Mismatches 56; Indels 2; Gaps 2;
Matches 330; Conservative 40; Mismatches 56; Indels 2; Gaps 2;

QY 7 YKVADIGLAAGRKALDIAENEMPGMLMRMRYSASKPLKGLARIAGCLHMTVETAVLIET 66
DB 6 YKVADISLAAGRKELAEIAENEMPGMLIAVREYAKSQPLKGLARIAGCLHMTVETAVLIET 65

QY 67 LVTLAGAEVQWSSCNIPSTQNHAAAAIAKAGIPYAWKGTDBEYLWCIEQTL-YPKDG-P 124
DB 66 LVALGAEVTVWSSCNISTYQDHAAAAIAATGVPVFAWKGTETBEYLWCIEQQLKSPSGKP 125

QY 125 LNMILDDGDLNLHTKYVOLLPGIRGISEETTTGVHNLKGMANGILKVPAINVDSV 184
DB 126 LNMILDDGDLNLVHERHPELVLDIRGISEETTTGVHNLKGMANGILKVPAINVDSV 185

QY 185 TSKFEDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVITETI 244
DB 186 TSKFEDNLGCRSLIDGILKRAVDVMIAGKAVVAGYGVGKCAQALRGFGARVITETI 245

QY 245 DPINALQAAMEGYEVTTMDACQEGNIFVTTGCIIDILGRHFEQMKDDAIVCNIGHFDV 304
DB 246 DPINALQAAMDGEVTTMTBEAVKEGQIFVTTGCRDIIRGEHFNEMKEDSIVCNIGHFDV 305

QY 305 EIDVWKLNAENAVEKNIKPOVDYRLKNGRRILLAEGLVNLGCMGHPSPVMSNFTN 364
DB 306 EIDVWKLNAENAVEKNIKPOVDYRLKNGRRILLAEGLVNLGCMGHPSPVMSNFTN 365

QY 365 QVMAQIELWTHPDKPVGVHFLPKLDEAVAHAHLKLVKLTKLTKTEKQAOYLGMSCDGP 424
DB 366 QVLAQIALWTDNTSYPLGVHMLPKLDEAVARHLKGLKGLVLTLSVQSDYLGIPVDGP 425

QY 425 FKPDHYRY 432
DB 426 YKADHYRY 433

RESULT 5
US-10-369-493-1740
Sequence 1740, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 1740
LENGTH: 449
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1740

Query Match 71.9%; Score 1633; DB 12; Length 449;
Best Local Similarity 70.8%; Pred. No. 3.7e-156; Mismatches 61; Indels 16; Gaps 3;
Matches 313; Conservative 52; Mismatches 61; Indels 16; Gaps 3;

QY 7 YKVADIGLAAGRKALDIAENEMPGMLMRMRYSASKPLKGLARIAGCLHMTVETAVLIET 66
DB 8 YKIADISLAAGRKELAEIAEHMPGLMAIRKAYGDVQPLKGLARIAGCLHMTVETAVLIET 67

QY 67 LVTLAGAEVQWSSCNIPSTQNHAAAAIAKAGIPYAWKGTDBEYLWCIEQTL-YPKDG-P 124
DB 68 LVALGAEVTVWSSCNISTYQDHAAAAIAASGVPVFAWKGTETBEYLWCIEQQLKSPSGKP 127

QY 125 LNMILDDGDLNLHTKYVOLLPGIRGISEETTTGVHNLKGMANGILKVPAINVDSV 184
DB 128 LNMILDDGDLNLVHERHPELVLDIRGISEETTTGVHNLKGMANGILKVPAINVDSV 187

QY 185 TSKFEDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVITETI 244
DB 188 TSKFEDNLGCRSLIDGILKRAVDVMIAGKAVVAGYGVGKCAQALRGFGARVITETI 247

QY 245 DPINALQAAMEGYEVTTMDACQEGNIFVTTGCIIDILGRHFEQMKDDAIVCNIGHFDV 304
DB 248 DPINALQAAMEGYVTTMEDASHIQGVFVTTTGCRTDIIRGEHFNEMKEDSIVCNIGHFDV 307

QY 305 EIDVWKLNAENAVEKNIKPOVDYRLKNGRRILLAEGLVNLGCMGHPSPVMSNFTN 364
DB 308 EIDVWKLNAENAVEKNIKPOVDYRLKNGRRILLAEGLVNLGCMGHPSPVMSNFTN 367

QY 365 QVMAQIELWTHPDKPVGVHFLPKLDEAVAHAHLKLVKLTKLTKTEKQAOYLGMSCDGP 410
DB 368 QVLAQIALWTDNTSYPLGVHMLPKLDEAVARHLKGLKGLVLTLSVQSDYLGIPVDGP 427

QY 411 EKQAOYLGMSCDGPFPKHRY 432
DB 428 KVQSEYLGIPBEGPPKADHYRY 449

RESULT 6
US-10-369-493-3683
Sequence 3683, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 3683
LENGTH: 447
TYPE: PRT
ORGANISM: Neurospora crassa
US-10-369-493-3683

Query Match 70.2%; Score 1594; DB 12; Length 447;
Best Local Similarity 69.8%; Pred. No. 3.2e-152; Mismatches 307; Conservative 52; Mismatches 65; Indels 16; Gaps 3;

Qy	7	YKVADI GLA WGRKALDI AENEM PGLM RRYE SAKPLKGARI AGCLHMTVETAVLIET	66
Db	8	FKVADLSLA AFGFKEI ELAENEM PGLM ATRK YIAADQPLKGARI AGCLHMTIQTAVLIET	67
Qy	67	LVTLGAEVQWSSCNI FSTONHAAAAIA KAGI PVYAWKGETDEEYLMCIROTLY - FKDG - P	124
Db	68	LTALGAETVWSSCNI FSTQDHAAAAIAAAGVPFAWKGETEBEYQWCLEQLLAFKDNKK	127
Qy	125	LNMLDDGGDLTNLIHTKYPOLLPIRGISEBT TTGVHNYLKMANGILKVPAINVNDSV	184
Db	128	LNMLDDGGDLTHLVHTKYPEMLEDCFGVSBETT TGVHLYRMLKEGKLLVPAINVNDSV	187
Qy	185	TKSKEDNLYGCRSLIDGIKRATDVMTAGVAVVAGYGVKGCAQALRGFGARVITEL	244
Db	189	TKSKEDNLYGCRSLVLDGIIKRAATDVMTAGIATVAVGFGDVKGCAALSGMARGARVITVE	247
Qy	245	DPINALQAAMBEYEVTTMDEACQEGNFVT TTGCDIILGRHFQMKDDAIVCNI GHFV	304
Db	248	DPINALQAAMAGYQVTTMEKAAPLQGFVTT TGCDDILVGKHFVEMDAIVCNI GHFV	307
Qy	305	EIDVQWLNENAVEKNIKPQVDRVRLKNGRRI ILLAEGRLVNI GCZMGHPSPFVMSNFTN	364
Db	308	EIDVAWLKANAASVQNIKPQVDRFLMKNGRHI ILLAEGRLVNI GCZMGHSPFVMSCSFTN	367
Qy	365	QVMAQIELLWTHPD-----KY PVGHFLPKKLDEAVAAHLGKLNVLTKLT	410
Db	368	QVLAQIMLYKANDEAFSNKYVEFGSKGLK EKVYVLPKILDEEVARILHDHCNVELTQUS	427
Qy	411	EKQAOYLGMSCDGP KPDPHY	430
Db	428	DVQAEYLG LATEGPYKSDQY	447

RESULT 7

US-10-369-493-10135
 ; Sequence 10135, Application US/10369493
 ; Publication NO. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 10135
 ; LENGTH: 426
 ; TYPE: PRT
 ; ORGANISM: magnetite-containing magnetic coccus
 US-10-369-493-10135

[illegible]

RESULT 8

```

RESULTS 8
US-10-369-493-13111
; Sequence 13111, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052) B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13111
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(516)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13111

```

Query Match	66.7%;	Score 1514.5;	DB 12;	Length 516;
Beat Local Similarity	60.5%;	pred. No. 4.4e-144;		
Matches 308;	Conservative 48;	Mismatches 70;	Indels 83;	Gaps 7;
Qy	7	YKVADI-----GLAAMGRKALDIAENMPGLMRMRERYSASKPLKGA	48	
Dbb	8	FKVADIVRLSTQLVEGEYHVXLESLAAGRREIELABIEPGLMAIRAKYGPAPQLKGA	67	
Qy	49	RIAGCLHMTV-----ETAVLIETLVTLGAEVQWSSCNIFSTQWHAAAA	91	
Dbb	68	RIAGCLHMTFRIPXELMNLNLTLPQFPQFVAVLIETLTALGAETVWTSNIFSTQDHAANA	127	
Qy	92	IAKAGIIPVYA-----WKGETDBEYLVWCIEQTL-YPKDG-PLNMLILDGG	133	
Dbb	128	IAAGVFPVYVDSQSYMSTKLIXAASPWKGETBEEYNWCLBQQLSAFKDQGLNLIILDGG	187	
Qy	134	DLTWLHHTKYPQLLPGIRISEETTGVGNLVKYMANGILKVPAINVNDSTVTKSKFDNLY	193	
Dbb	188	DLTSLVHTKYPEQLQCYGLSEETTGTVHHVLYMLKBGKLLVPAINVNDSTVTKSKFDNLY	247	
Qy	194	GCRESLIDGIKRATDVIYMIAGKAVAVAGYGDVGKCAQALRFGFARVITEIDPINALQAA	253	
Dbb	248	GCRESLIDGIKRATDVIYMIAGKAVAVAGYGDVGKCAQALHTWVGARVITEIDPINALQAA	307	

Query Match	64.5%	Score 1465;	DB 12;	Length 428;
Best Local Similarity	66.2%;	Pred. No. 3.3e-139;		
Matches	284;	Conservative	58;	Mismatches 83; Indels 4; Gaps 3;
Qy	7	YKVADIGLAWGKALDIAENEMPLMRBRYASAKPLKAGRTAGCLHMTVETAVLIET	66	
Db	1	HKVNDISLAWGRIEILAEAEMLMSIRKEFGPSKPLKRGAVAGCLHMTIQIQTAVLIET	60	
Qy	67	LVTLGAEVQSSNCNIFSTQNHAAAIARAKGIPVYAWKGETDEEYDLWCIEQTLFYKDG--P	124	
Db	61	LIELGAEVWSSNCNIFSTQDHAIAAAGISVYAWKGMNEEFPDWCIEQTLFFGEDRKP	120	
Qy	125	LNMLIDGGDLTNLIHFKYPLLPGIRGISSETTTGVHNLKMMANGILKVPAINVDSV	184	
Db	121	LNMLIDGGDLTNMVLDFPELVKDIRGISSETTTGVLRKDRBRNGSLVLPAINVDSV	180	
Qy	185	TKSKFDNLYGRESLIDGIKRATDMVITAGKVAVVAGYGDVGKCAQALRGFGARVITEI	244	
Db	181	TKSKFDNKYCKESLVDSSIRRAITDMMAGKVAVVAGYGDVGKGSAAILRGAGARVITEI	240	
Qy	245	DPINALQAAMEGYEVTWMDACQSGNTFVTTTGCIDILGRHFQMKDDAIVCNIGHFDV	304	
Db	241	DPICALQAAMDGYEVKKQADAVKRAIDVVTATGKNKIITGEHFFKAMRDKVIVCNIGHFDN	300	
Qy	305	EIDVKVLNEN-AVEKVNIKPOVDYRLKNGRRRIILLAEGRVLNVLGCAMGHPFSFVMSNST	363	
Db	301	EIDMAVLNKTYGSKVTYVKPQVDIYNV-DGHVDVILLAEGRVLNVLGCATGHPFSFVMSSEFS	359	
Qy	364	NOVMAQIELMTHDPKYFVGVHFLPKKLDEAVAEAHGLKGLNKLTKLITEKQAOYLGWCDG	423	
Db	360	NOVIAQIELMENSCKYENKVYTLPKSLDEKVARLHLKSIDVELDILSADOAAYIGTVVDG	419	

Query Match	62.9%;	Score 1428;	DB 12;	Length 468;
Best Local Similarity	61.3%;	Pred. No. 2.1e-135;		
Matches	287;	Conservative 52;	Mismatches 87;	Indels 42; Gaps 6;
Qy	7	YKVADIGLAANGRKALDIAENEMPGLMRREYSASKPLKGARIAACCLHMTVETAVLIET	66	
Db	1	YKIADIISLADWGRKEIDIAEHMPGLMIRRYASKQPLKGVVRVTGSLHMTIQTAVLIET	60	
Qy	67	LVTIGAEEVOWSSCNIFSTQNHAARAAIAKAGIPVAWKGETDEYLWCIEOTLYP--KDGP	124	
Db	61	LKDIGADVRAWASCNIFSTDHAAAAAIATSGTPPAWKGETLEBYDWCTTQALTFTTLADGT	120	
Qy	125	L---NMILDDGGDTNLTHTKY-----PQLT-----PGI---	150	
Db	121	LTGPBLIVDDCGDATLLHKGYELENGSTWVDPSDLSEEQVIKRLLKRIAIERPGYWTR	180	
Qy	151	----RGISETTTTGVNHLKYOMANGILLKVPAINVNDSVTSKFDNLYGCRESLIDGIKR	205	
Db	181	VNDWKGVSBETTTGVHRLYQIAATGRLLVPAINVNDSVTSKFDNLYGCRESLADGLKR	240	
Qy	206	ATDVMIAAGKAVVAGYGVDVGKCAQALRFGGARVIIETEIDPINALOANMEGYEVTIMDEA	265	
Db	241	AMDVMLAGKLAVCCGYDVGKGSASHRAYGARVITVEDIPTCALOANMEGPEVTTVEDT	300	
Qy	266	CQEGNI FVTYTTCGTDII LGRHFEOMKDDAICVNICHPDVEIDVKWLNE--NAVEKNVIKPQ	324	
Db	301	LGQADI VTTTGKDVIRIEHTAMQDVIVCNI GHFPDNEIQUDALNTLTGVOKINIKPQ	360	
Qy	325	VDYRLXKNRRRI ILLAEGRVLNLCGAMGHSPFVMSNSFTNQVMAQTIELMTHPDKPYGVGH	384	
Db	361	VDKFILPNGNTLFLLAEGRVLNLCGCATGHSPFVMSNSFANQTLAQIDLGNQKDVYEKNVY	420	
Qy	385	FLPKKLDEAVAHLGKUNLVKTUTEKOQAQYLGMSCDGPFPRDHTRY	432	
Db	421	RLPKKLDEAVARLHEKIGVKKLTITANQAAYLGISVEGPPKPEHYRY	468	

RESULT 11 .
US-10-369-493-9386
; Sequence 9386, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 9386
LENGTH: 468
TYPE: PRT
ORGANISM: Xylella fastidiosa
US-10-369-493-9386

Query Match 62.6%; Score 1422; DB 12; Length 468;
Best Local Similarity 61.1%; Pred. No. 8.6e-135;
Matches 286; Conservative 52; Mismatches 88; Indels 42; Gaps 6;

QY 7 YKVDIGLAAGWKALDIAENEMPGMLMRMRYSASKPLKGARIAGCLHMTVETAVLIET 66
DB 1 YKIADISLADWGRKEIDIAEHMPGLMSIRRYKVASQPLKGVRTGSLHMTVETAVLIET 60
QY 67 LVTLGAEVOWSSCNIFSTQNHAAAAIAKAGIPYANKGETDEEYLWCIEQTLYF--KQGP 124
DB 61 LKDIGADVWASCNIFSTQDHAAAAIATSGTPVFAWKGETLEEYDCTIQALTFTLADGT 120
QY 125 L---NMILDDGGDLTLNHTKY-----PQL-----PGI--- 150
DB 121 LTPGLIVDDGGDATTLLHKGYELENGSTWDEPSLSSEQVTKRLKLLKIAIERPGYWR 180
QY 151 -----RGISETTGVHNLKYMANGILKVPAINVNDSTVTKSFNDLYGCRSLIDGIKR 205
DB 181 VNDWKGVSSEETTTGVHRLYQIAATGRLLVPAINVNDSTVTKSFNDLYGCRSLADGLKR 240
QY 206 ATDVMTAGVAVVAGYGVGKCAQALRGFGARVITTEIDPINALQAMGEGEVTTMDRA 265
DB 241 AMDVMTAGVAVVAGYGVGKCAQALRGFGARVITTEIDPINALQAMGEGEVTTMDRA 300
QY 266 COEGNIFVTTTGCIDIIILGRHFQMKDDAIVCNIGHFDEIDVKWLN--NAVEKNIKQ 324
DB 301 LGGADIVTTTGNKQVIRIEHTMTAKDQVIVCNIGHFDEIDVKWLN--NAVEKNIKQ 360
QY 325 VDRYLRKNGRRRIILLAEGRVLNLCAMGHPSPFVMSNFTNQVMAQIELWTHPDKYPGVH 384
DB 361 VDKFILPNGNTLLAEGRVLNLCAMGHPSPFVMSNFTNQVMAQIELWTHPDKYPGVH 420
QY 385 FLPKLDEAVAHAHLKLVNLTKEQAQYLGMSCDGPFKPDHYRY 432
DB 421 RUPKLDDEEARLHLEKIGVLTTLTANQAAYLGISVEGPFKPDHYRY 468

RESULT 12
US-10-369-493-9266
Sequence 8266, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374

Query Match 61.9%; Score 1405; DB 12; Length 466;
Best Local Similarity 60.9%; Pred. No. 4.5e-133;
Matches 287; Conservative 53; Mismatches 83; Indels 48; Gaps 7;

QY 4 KLPYKVADIGLAAGWKALDIAENEMPGMLMRMRYSASKPLKGARIAGCLHMTVETAVL 63

SEQ ID NO 8266
LENGTH: 476
TYPE: PRT
ORGANISM: Thermobifida fusca
US-10-369-493-8266

Query Match 61.9%; Score 1406; DB 12; Length 476;
Best Local Similarity 58.9%; Pred. No. 3.7e-133;
Matches 279; Conservative 57; Mismatches 90; Indels 48; Gaps 4;

QY 7 YKVDIGLAAGWKALDIAENEMPGMLMRMRYSASKPLKGARIAGCLHMTVETAVLIET 66
DB 3 FKVADLSLAEPGKKEIRLAHEMPGLMATRASFGPKPLRGAKITGSLHMTVETAVLIET 62
QY 67 LVTLGAEVOWSSCNIFSTQNHAAAAIA-----KAGIPYANKGETDEEYLWCIEQT 117
DB 63 LVELGAEVWASCNIFSTQDHAAAAAVVVGPDGTPDNPRGVPVFAWKGETLEEYVWCTEQ 122
QY 118 LYFKDGP-LNMILDDGGDLTLNLIH----- 140
DB 123 LTPGEGENMILDDGGDATMLVHKGVQYKAGAVDPDPTADSEPRIVLELLRRLAEN 182
QY 141 -TKYPQLPFGIRGISETTGVHNLKYMANGILKVPAINVNDSTVTKSFNDLYGCRSL 199
DB 183 PRKWTIASRIKGVTEETTTGVHRLYEMQSGTLLFPAINVNDSTVTKSFNDKNGYGRHSL 242
QY 200 IDGIKRAVTMTAGVAVVAGYGVGKCAQALRGFGARVITTEIDPINALQAMGEGEV 259
DB 243 IDGINRATDVLTGGKAVVAGYGVGKCAESLRGQGARVITTEIDPICALQAMDGQV 302
QY 260 TTMDEACOBGNIFVTTTGCIDIIILGRHFQMKDDAIVCNIGHFDEIDVKWLN--NAVEK 318
DB 303 TTLDEWETADIFATATGNRQVITAEHARMKQQAIVGNIGHFDEIDWAGLAATPGIEK 362
QY 319 VNKPQVDRYLRKNGRRRIILLAEGRVLNLCAMGHPSPFVMSNFTNQVMAQIELWTHPK 378
DB 363 IEIKPQVHWRFPDGHISIVLSEGRLLNLGNATGHPSPFVMSNFTNQVMAQIELWTHPK 422
QY 379 YPVGHFLPKLDEAVAHAHLKLVNLTKEQAQYLGMSCDGPFKPDHYRY 432
DB 423 YPTGVYLPKHLDEKVARLHLDALGVKLTKEQAQYLGMSCDGPFKPDHYRY 476

RESULT 13
US-10-369-493-8750
Sequence 8750, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 8750
LENGTH: 466
TYPE: PRT
ORGANISM: Ralstonia metallidurans
US-10-369-493-8750

Query Match 61.9%; Score 1405; DB 12; Length 466;
Best Local Similarity 60.9%; Pred. No. 4.5e-133;
Matches 287; Conservative 53; Mismatches 83; Indels 48; Gaps 7;

QY 4 KLPYKVADIGLAAGWKALDIAENEMPGMLMRMRYSASKPLKGARIAGCLHMTVETAVL 63

Db 2 KQDYIVADIGLAGWGRKKEIAIAETEMPGMLMAIRDEFADAAAQPLKGARIAGSLHMTIQTAVL 61
QY 64 IETVLTLGAEVQWSSCNIFSTQNHAAAAIAKAGIPYVANKGETDEBYLWCIEQTLX-FKD 122
Db 62 IETLKALGADVWASCNIFSTQDHAAAAIAASGTPVAFKGSLSKEY-WDFTHRIIFDWD 120
QY 123 GPL-NMILDDGGDLTNLIH-----TKYPOL 146
Db 121 GGTNNMILDDGGDATALLLHLGARAEDASLIAKPTSEETFLFAAIKEKLAODSTWYSRN 180
QY 147 LPIRGISSETTGVHNLKYMANGILKVPAINVNDVTSKSPDNLYGCRSLIDGIKRA 206
Db 181 LAAIRGVTEETGGVHRLYQMAQKGLKFPAINVNDVTSKSPDNLYGCRSLVDGIKRA 240
QY 207 TDVMIAGKAVVAGYGVGKGAQAALRGFARVIIITEIDPINALQAAMEGYEYVTTMDEAC 266
Db 241 TDVMIAGKAVVAGYGVGKGAQAALRALSAQVWVTEIDPICALQAAMEGYRVVTDYAA 300
QY 267 QEGNIFVTTTCIDIIILGRHFEQMKDDAIVCNIGHFDEVIDVKMLNENAVEKY---NIXP 323
Db 301 EHGDIPTCTGNVHTHDMAKMKDQAIVCNIGHFDEVIDIA-----SVEKYEWDEIKP 355
QY 324 QVDYRLKNGRRILLAEGRNLNLCAMGHPSPVMSNFTNQVMAQIELWTHPD--KYPV 381
Db 356 QVDHVPDPGKKIILAKRGLNLCATGHPSPVMSNFTNQVMAQIELWTHPD--KYPV 415
QY 382 GVHFLPKLDEAAVAEHLGKLNKLTETEQAQOYLGMSCDGPFPKPDHYRY 432
Db 416 GVTVLPKHLDEKVARLQRLKLNQLTELTEQQAAYIGVKEGYPKADHYRY 466

RESULT 14
US-10-369-493-15745
; Sequence 15745, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15745
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15745

Query Match 61.7%; Score 1401; DB 12; Length 469;
Best Local Similarity 60.9%; Pred. No. 1.1e-132;
Matches 285; Conservative 45; Mismatches 96; Indels 42; Gaps 6;

QY 7 YKVADIGLAAGWKALDIAENMPGLMRMRERYASASKPLKGARIAGCLHMTVETAVLIET 66
Db 2 YKIADISLADWGRKELDIAEHMPGLMSIRRKHAQTKPLKDVRIITGSLHMTIQTAVLIET 61
QY 67 LVTLGAEVQWSSCNIFSTQNHAAAAIAKAGIPYVANKGETDEBYLWCIEQTLX-FKDGP 124
Db 62 LKDIGANVRWASCNIFSTQDHAAAAIAATGTPVFAWKGETLEBYWDCDTLDTLFTLPDGT 121
QY 125 L---NMILDDGGDLTNLIHXY-----POLL-----PGI--- 150
Db 122 LTGPELVDDGGDVTLLIHKGYLENGSTWVDEPASPHEEGVTKALKRVAVERPGYWAR 181
QY 151 -----RGISEETTTGVHNLKYMANGILKVPAINVNDVTSKSPDNLYGCRSLIDGIK 205
Db 2 YKIADISLADWGRKELDIAEHMPGLMSIRRKHAQTKPLKDVRIITGSLHMTIQTAVLIET 61
QY 67 LVTLGAEVQWSSCNIFSTQNHAAAAIAKAGIPYVANKGETDEBYLWCIEQTLX-FKDGP 124
Db 62 LKDIGANVRWASCNIFSTQDHAAAAIAATGTPVFAWKGETLEBYWDCDTLDTLFTLPDGT 121
QY 125 L---NMILDDGGDLTNLIHXY-----POLL-----PGI--- 150
Db 122 LTGPELVDDGGDVTLLIHKGYLENGSTWVDEPASPHEEGVTKALKRVAVERPGYWAR 181
QY 151 -----RGISEETTTGVHNLKYMANGILKVPAINVNDVTSKSPDNLYGCRSLIDGIK 205

Db 182 VVKDWKGVSEETTTGVHRLYQIAEAGKLLIPAINVNDVTSKSPDNLYGCRSLADGLKR 241
QY 206 ATDVMIAGKAVVAGYGVGKGAQAALRGFARVIIITEIDPINALQAAMEGYEYVTTMDEA 265
Db 242 AMDVNLAKVAVVCGYGVGKGAQAALRGFARVIIITEIDPICALQAAMEGYEYVTTIEST 301
QY 266 COEGNIFVTTTCIDIIILGRHFEQMKDDAIVCNIGHFDEVIDVKWLNE-NAVEKVNKIQP 324
Db 302 LGRGDIYVTTTGNKDIITVEHLQAMKDDAIVCNIGHFDEVIDIQLDLWEKEDSVEKKYV 361
QY 325 VDRYRLKNGRRILLAEGRNLNLCAMGHPSPVMSNFTNQVMAQIELWTHPDKYPVGVH 384
Db 362 VDKYVFGNGAIFLLADGRLNLCATGHPSPVMSNFTNQVMAQIELWTHPDKYPVGVH 421
QY 385 FUPKKLDEAAVAEHLGKLNKLTETEQAQOYLGMSCDGPFPKPDHYRY 432
Db 422 ILPKHLDEAAVAEHLGKLNKLTETEQAQOYLGMSCDGPFPKPDHYRY 469

RESULT 15
US-10-369-493-16129
; Sequence 16129, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16129
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-16129

Query Match 61.7%; Score 1401; DB 12; Length 469;
Best Local Similarity 60.9%; Pred. No. 1.1e-132;
Matches 285; Conservative 45; Mismatches 96; Indels 42; Gaps 6;

QY 7 YKVADIGLAAGWKALDIAENMPGLMRMRERYASASKPLKGARIAGCLHMTVETAVLIET 66
Db 2 YKIADISLADWGRKELDIAEHMPGLMSIRRKHAQTKPLKDVRIITGSLHMTIQTAVLIET 61
QY 67 LVTLGAEVQWSSCNIFSTQNHAAAAIAKAGIPYVANKGETDEBYLWCIEQTLX-FKDGP 124
Db 62 LKDIGANVRWASCNIFSTQDHAAAAIAATGTPVFAWKGETLEBYWDCDTLDTLFTLPDGT 121
QY 125 L---NMILDDGGDLTNLIHXY-----POLL-----PGI--- 150
Db 122 LTGPELVDDGGDVTLLIHKGYLENGSTWVDEPASPHEEGVTKALKRVAVERPGYWAR 181
QY 151 -----RGISEETTTGVHNLKYMANGILKVPAINVNDVTSKSPDNLYGCRSLIDGIK 205
Db 182 VVKDWKGVSEETTTGVHRLYQIAEAGKLLIPAINVNDVTSKSPDNLYGCRSLADGLKR 241
QY 206 ATDVMIAGKAVVAGYGVGKGAQAALRGFARVIIITEIDPINALQAAMEGYEYVTTMDEA 265
Db 242 AMDVNLAKVAVVCGYGVGKGAQAALRGFARVIIITEIDPICALQAAMEGYEYVTTIEST 301
QY 266 COEGNIFVTTTCIDIIILGRHFEQMKDDAIVCNIGHFDEVIDVKWLNE-NAVEKVNKIQP 324
Db 302 LGRGDIYVTTTGNKDIITVEHLQAMKDDAIVCNIGHFDEVIDIQLDLWEKEDSVEKKYV 361
QY 325 VDRYRLKNGRRILLAEGRNLNLCAMGHPSPVMSNFTNQVMAQIELWTHPDKYPVGVH 384

Db 362 VDKYVFGNGNAIFLLADGRVLNLGCATGHPSPVMSNSFANQTLAQIDLMEKRDSEYKQVY 421
Qy 385 FLPKLDEAVAEAHILGKLVKLTLEKQAOYLGMSCDGPFPKDPHYRY 432
Db 422 ILPKHLDDEVARLHLEKIGVKLTTLTKOQADYLGVDVAGPYKDPHYRY 469

Search completed: December 30, 2003, 10:47:06
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2003, 10:39:38 ; Search time 21 Seconds
(without alignments)
870.395 Million cell updates/sec

Title: US-10-043-787-1

Perfect score: 2271

Sequence: 1 MSDKLPYKVADIGLAAGRK.....QAQLGMSCDGPFKPDHYRY 432

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2271	100.0	432	2	US-08-896-005-4
2	2271	100.0	432	4	US-09-347-878-1
3	2193	96.6	432	2	US-08-896-005-5
4	1289	56.8	485	3	US-08-930-894-2
5	1249.5	55.0	502	4	US-09-252-991A-23501
6	1222	53.8	463	4	US-09-328-352-4600
7	1149.5	50.6	504	2	US-08-896-005-3
8	1146	50.5	500	2	US-08-896-005-1
9	421	18.5	97	1	US-08-204-740-7
10	421	18.5	97	3	US-09-081-167A-7
11	421	18.5	97	3	US-09-081-395-7
12	421	18.5	97	3	US-09-416-833-7
13	421	18.5	97	5	PCT-US95-02521-7
14	388	17.1	95	1	US-08-204-740-9
15	388	17.1	95	3	US-09-081-167A-9
16	388	17.1	95	3	US-09-081-395-9
17	388	17.1	95	3	US-09-416-833-9
18	388	17.1	95	5	PCT-US95-02521-9
19	347	15.3	99	3	US-08-930-894-8
20	249.5	11.0	138	3	US-08-930-894-5
21	229.5	10.1	140	3	US-08-930-894-4
22	228	10.0	139	3	US-08-930-894-7
23	224	9.9	139	3	US-08-930-894-6
24	112.5	5.0	485	4	US-09-252-991A-25242
25	103.5	4.6	257	4	US-09-134-001C-3562
26	98.5	4.3	683	4	US-09-816-093-2
27	98	4.3	449	4	US-09-530-836-2

28	98	4.3	1302	1	US-08-232-537-2	Sequence 2, Appl1
29	461	4.3	461	4	US-09-134-001C-3892	Sequence 3892, Ap
30	97.5	4.3	539	4	US-09-134-001C-5118	Sequence 5118, Ap
31	96.5	4.2	1261	3	US-09-080-855-2	Sequence 2, Appl1
32	96.5	4.2	1261	4	US-09-566-076-2	Sequence 2, Appl1
33	95.5	4.2	349	4	US-09-134-001C-4884	Sequence 4884, Ap
34	93.5	4.1	345	3	US-09-222-817-2	Sequence 2, Appl1
35	93.5	4.1	345	3	US-09-222-786-2	Sequence 12, Appl1
36	93.5	4.1	530	3	US-09-222-817-12	Sequence 12, Appl1
37	93.5	4.1	530	3	US-09-222-817-14	Sequence 14, Appl1
38	93.5	4.1	530	3	US-09-222-786-12	Sequence 12, Appl1
39	93.5	4.1	530	3	US-09-222-786-14	Sequence 14, Appl1
40	93	4.1	480	4	US-09-252-991A-28668	Sequence 28668, A
41	92	4.1	586	4	US-09-134-001C-4456	Sequence 4456, Ap
42	92	4.1	1297	4	US-09-328-352-6373	Sequence 6373, Ap
43	91	4.0	546	3	US-08-470-260-6	Sequence 6, Appl1
44	91	4.0	546	3	US-08-471-491-6	Sequence 6, Appl1
45	91	4.0	546	3	US-08-466-662-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-08-896-005-4
; Sequence 4, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,005
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0337 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 178277
; US-08-896-005-4

Query Match 100.0%; Score 2271; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.2e-246;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASGSKPLKGARIAGCLHMTVET 60
DB 1 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASGSKPLKGARIAGCLHMTVET 60
QY 61 AVLIETLVTLAGVQWSSCNIFSTQNHAAIAAKAGIPVYAWKGETDEEYLCWICIEQTLF 120
DB 61 AVLIETLVTLAGVQWSSCNIFSTQNHAAIAAKAGIPVYAWKGETDEEYLCWICIEQTLF 120
QY 121 KQGPLNMILDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLKYMANGILKVPAINV 180
DB 121 KQGPLNMILDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLKYMANGILKVPAINV 180
QY 181 NDSVTSKFDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
DB 181 NDSVTSKFDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
QY 241 ITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIG 300
DB 241 ITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIG 300
QY 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFWSN 360
DB 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFWSN 360
QY 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKLDEAVABHLGKLNKLTTEKQAOYLGM 420
DB 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKLDEAVABHLGKLNKLTTEKQAOYLGM 420
QY 421 CDGPFKPDHYRY 432
DB 421 CDGPFKPDHYRY 432
RESULT 2
US-09-347-878-1
; Sequence 1, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human S-adenosylhomocysteine hydrolase protein
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M61831/GenBank
US-09-347-878-1
Query Match 100.0%; Score 2271; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.2e-246;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASGSKPLKGARIAGCLHMTVET 60
DB 1 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASGSKPLKGARIAGCLHMTVET 60
QY 61 AVLIETLVTLAGVQWSSCNIFSTQNHAAIAAKAGIPVYAWKGETDEEYLCWICIEQTLF 120
DB 61 AVLIETLVTLAGVQWSSCNIFSTQNHAAIAAKAGIPVYAWKGETDEEYLCWICIEQTLF 120
QY 121 KQGPLNMILDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLKYMANGILKVPAINV 180
DB 121 KQGPLNMILDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVHNLKYMANGILKVPAINV 180

QY 181 NDSVTSKFDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
DB 181 NDSVTSKFDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
QY 241 ITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIG 300
DB 241 ITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIG 300
QY 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFWSN 360
DB 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFWSN 360
QY 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKLDEAVABHLGKLNKLTTEKQAOYLGM 420
DB 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKLDEAVABHLGKLNKLTTEKQAOYLGM 420
QY 421 CDGPFKPDHYRY 432
DB 421 CDGPFKPDHYRY 432
RESULT 3
US-08-896-005-5
; Sequence 5, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,005
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0337 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 904132
US-08-896-005-5
Query Match 96.6%; Score 2193; DB 2; Length 432;
Best Local Similarity 96.5%; Pred. No. 7.1e-238;
Matches 417; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
QY 1 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASGSKPLKGARIAGCLHMTVET 60

Db 1 MSCLKPYKADIGLAAGRKALDIAENEMPGLMRREMTSASKPLKAGARIAGCLHMTVET 60
Qy 61 AVLIETLVTLAGAEVOWSSCNIFSTONHAAATAKAGIPVYANKGETDEEYLWCIEOTLYF 120
Db 61 AVLIETLVALGAEVWRSSCNIFSTODHAAATAKAGIPVYANKGETDEEYLWCIEOTLYF 120
Qy 121 KDGPLNMLDDGGDLTNLHTKYPOLLPGRIGISEBTITGVNLYKQAMANGILKVPAINV 180
Db 121 KDGPLNMLDDGGDLTNLHTKYPOLLSGIRIGISEBTITGVNLYKQAMANGILKVPAINV 180
Qy 181 NDSVTKSFENLYGCRRESIDIGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
Db 181 NDSVTKSFENLYGCRRESIDIGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
Qy 241 ITEIDPINALQAAAMEGYEVTWDEACQEGNIFVTTTGCIDIIILGRHFEQMKODAIVCNIG 300
Db 241 ITEIDPINALQAAAMEGYEVTWDEACQEGNIFVTTTGCIDIIILGRHFEQMKODAIVCNIG 300
Qy 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGRVLNLCAMGHPSPFWMGN 360
Db 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGRVLNLCAMGHPSPFWMGN 360
Qy 361 SFTNQVMAQIELMTHDPKYPVGVHFLPKKLDEAVAEHLGKLVNKLTKLTKTEKQAOYLGM 420
Db 361 SFTNQVMAQIELMTHDPKYPVGVHFLPKKLDEAVAEHLGKLVNKLTKLTKTEKQAOYLGM 420
Qy 421 CDGPKPKPDHYRY 432
Db 421 INGPKPKPDHYRY 432

RESULT 4
US-08-930-894-2
; Sequence 2, Application US/08930894
; Patent No. 6037524
; GENERAL INFORMATION:
; APPLICANT: GREENLAND, Andrew James
; APPLICANT: DRAPER, John
; APPLICANT: SKIPSEY, Marc
; APPLICANT: WARNER, Simon
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/930,894
; FILING DATE: 09-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00882
; FILING DATE: 10-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9507381.3
; FILING DATE: 10-APR-1995
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-930-894-2

Query Match 56.8%; Score 1289; DB 3; Length 485;
Best Local Similarity 56.1%; Pred. No. 5.1e-136;
Matches 265; Conservative 55; Mismatches 106; Indels 46; Gaps 5;
Qy 7 YKVADIGLAAGRKALDIAENEMPGLMRREMTSASKPLKAGARIAGCLHMTVETAVLIET 66
Db 14 YKVKDMSQADFQRLBIELAEVEMPGMACRAEFGPAQPFKAGKITGSLHMTTQTAVLIET 73
Qy 67 LVTLAGAEVOWSSCNIFSTONHAAATAKAGIPVYANKGETDEEYLWCIEOTLY-YPKDGSL 125
Db 74 LTALGAEVOWSSCNIFSTODHAAATAKAGIPVYANKGETLQEWCTERALDWDGPGGP 133
Qy 126 NMILDGGDLTNLH-----TKYP 144
Db 134 DLIVDDGGDTLLIHEGVKAESEYKTKMPDPAGTDNAEFOIVLTIIRDGLKVDPTKYR 193
Qy 145 QLLPGIRIGISEBTITGVNLYKQAMANGILKVPAINVDSVTKSKEDNLYGCRRESIDIGIK 204
Db 194 KMKDRIVGVSEETTTGVKRLYQMANNSLLFPAINVDSVTKSKEDNLYGCRRESIDIGIK 253
Qy 205 RATDVMIAGKAVVAGYGVGKCAQALRGFGARVIITEIDPINALQAAAMEGYEVTWDE 264
Db 254 RATDVMIAGKAVVAGYGVGKCAQALRGFGARVIITEIDPINALQAAAMEGYEVTWDE 313
Qy 265 AQCEGNIFVTTTGCIDIIILGRHFEQMKODAIVCNIGHFDEIDVKWLNENAVEKVNIPK 323
Db 314 VVSEADIFVTTTGNKDIIIMLDMRKNNAIVCNIGHFDEIDMLGLEYTPGKIKRITIKP 373
Qy 324 QVDYRL-KNGRRILLAEGRVLNLCAMGHPSPFWMNSFTNQVMAQIELMTH--PDKYP 380
Db 374 QTDWRVFPETGTIIVLAEGRLNLCATGHPSPFWMNSFTNQVMAQIELMTH--PDKYP 433
Qy 381 VGVHFLPKKLDEAVAEHLGKLVNKLTKLTKTEKQAOYLGMSCDGPCKPDHYRY 432
Db 434 KKYTVLPKHLDEKVAALHLGKLVNKLTKLTKLSPSQADYISVPIEGPKPPHYRY 485

RESULT 5
US-09-252-991A-23501
; Sequence 23501, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23501
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23501

Query Match 55.0%; Score 1249.5; DB 4; Length 502;
Best Local Similarity 56.5%; Pred. No. 1.5e-131;
Matches 261; Conservative 59; Mismatches 101; Indels 41; Gaps 9;
Qy 7 YKVADIGLAAGRKALDIAENEMPGLMRREMTSASKPLKAGARIAGCLHMTVETAVLIET 66
Db 46 YKVADITLAAGRRRLIIAEESEMPALMURRYAGQQPLKAGKITGSLHMTTQTAVLIET 105
Qy 67 LVTLAGAEVOWSSCNIFSTONHAAATAKAGIPVYANKGETDEEYLWCIEOTLYFKDG-- 123
Db 106 LVALGAEVOWSSCNIFSTODHAAATAKAGIPVYANKGETEYEWICIEQTI-LKDGQPW 164
Qy 124 PLNMILDGGDLTNLHTKYPOLLPGRIGISEBTITGVNLYKQAMANGILKVPAINVDS 183

Db 165 DANVLDDGDLTEILHKKYQPMLEIRHIGITEETTTGVHRLDMLKNGALKVPAINVND 224
Qy 184 VTKSKDNLYGCRSLIDGKTRATDVMIAGKAVVAGYGVGKCAQALRGFCARVITTE 243
Db 225 VTKSKDNKYGCRSLIDGKTRATDVMIAGKAVVAGYGVGKCAQALRGFCARVITTE 284
Qy 244 IDPINALQAAMEGYEVTT-MDEACQEG-----NIFVTTTGCIDIIILGRHFEQ 289
Db 285 VDPICAMQACMDGFEVSPYKNGINDGTEASIDAALLGKIDLIIVTTGNVNCVNDANMLKA 344
Qy 290 MKDDAIVCNIGHDFEIDVVKMLNEN-AVEKNVTKPV-----DRYRLKNGRRILLIA 340
Db 345 LKRAVVCNIGHDFEIDVVKMLNEN-AVEKNVTKPV-----DRYRLKNGRRILLIA 402
Qy 341 EGRVLNLCAMGCHPSFVMSNFTNOVMAQIELWTHDPKYP-----VGHFPLPKL 390
Db 403 EGRVLNLCAMGCHPSFVMSNFTNOVMAQIELWTHDPKYP-----VGHFPLPKL 460
Qy 391 DEAVAHAHLGKLVNKLTKTEKQAQYLGMSCDGPFKPDHRY 432
Db 461 DEVALEMYGFGGVVTLTPKQAEVIGSVGEGPFKPDHRY 502

RESULT 6.

US-09-328-352-4600
; Sequence 4600, Application US/09328352
; Patent No. 6562958

GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 4600

; LENGTH: 463

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-4600

Query Match 53.8%; Score 1222; DB 4; Length 463;
Best Local Similarity 56.2%; Pred. No. 1.6e-128;
Matches 255; Conservative 62; Mismatches 105; Indels 32; Gaps 8;

Qy 7 YKVADIGLAWRKALDIAENEMPGLMRERYASPKLKGARIAGCLHMTVETAVLIET 66
Db 14 YKVADISLADYGRKEIKLAEPALIGRLKRYAASKPLAGAKILGCIHMTIQTAVLIET 73
Qy 67 LVTGAEVOWSSCNIFSTONHAAAAIAKAGIPVYAWKGETDEEYLWCIEOTLYFKDGP-- 124
Db 74 LVELGAEVWTSNIFSTONHAAAAIAAGIPVYAWKGETDEEYLWCIEOTLYFKDGP 133
Qy 125 LNMILDDGDLTLNLIHTKYPOLLPGIRGISEETTTGVHNLKYMANGILKVPAINVND 184
Db 134 ANMILDDGDLTLNLIHTKYPOLLPGIRGISEETTTGVHNLKYMANGILKVPAINVND 193
Qy 185 TSKKFNLYGCRSLIDGKTRATDVMIAGKAVVAGYGVGKCAQALRGFCARVITTE 244
Db 194 TSKKFNLYGCRSLIDGKTRATDVMIAGKAVVAGYGVGKCAQALRGFCARVITTE 253
Qy 245 DPINALQAAMEGYEVTT-MDEACQEG-----NIFVTTTGCIDIIILGRHFEQ 290
Db 254 DPCAMQACMDGFEVSPYKNGINDGTEASIDAALLGKIDLIIVTTGNVNCVNDANML 313
Qy 291 KDDAIVCNIGHDFEIDVVKMLNEN-AVEKNVTKPV-----DRYRLKNGRRILLIA 348
Db 314 KAGAVVCNIGHDFEIDVVKMLNEN-AVEKNVTKPV-----DRYRLKNGRRILLIA 371
Qy 349 CAMGCHPSFVMSNFTNOVMAQIELWTHDPKYP-----VGHFPLPKLDEAVAHAH 398
Db 372 NATGHPSPRVMSNFTNOVMAQIELWTHDPKYP-----VGHFPLPKLDEAVAHAH 429

Qy 399 LGLKLVNKLTKTEKQAQYLGMSCDGPFKPDHRY 432
Db 430 VAGFGVLTQLTQEQADYLGIAVEGPFKPSDAYKY 463

RESULT 7

US-08-896-005-3

; Sequence 3, Application US/08896005
; Patent No. 5854023

GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/896,005

; FILING DATE: Filed Herewith

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0337 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 504 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 969078

US-08-896-005-3

Query Match 50.6%; Score 1149.5; DB 2; Length 504;
Best Local Similarity 51.5%; Pred. No. 2.8e-120;
Matches 218; Conservative 83; Mismatches 121; Indels 1; Gaps 1;

Qy 9 VADIGLAWRKALDIAENEMPGLMRERYASPKLKGARIAGCLHMTVETAVLIETLV 68
Db 70 VKSISKSAFGRREIBIAESEMFGIMTKRANDEKPLKGANIVGCTHVNQAQSAVLIETLV 129
Qy 69 TLGAEVOWSSCNIFSTONHAAAAIAKAGIPVYAWKGETDEEYLWCIEOTLYFKDGPLNMI 128
Db 130 QLGAIVRWAACNISTONHAAAAIAAGIPVYAWKGETDEEYLWCIEOTLYFKDGPLNMI 189
Qy 129 LDDGGDLTLNLIHTKYPOLLPGIRGISEETTTGVHNLKYMANGILKVPAINVND 188
Db 190 LDDGGDLTLNLIHTKYPOLLPGIRGISEETTTGVHNLKYMANGILKVPAINVND 249
Qy 189 FDNLYGCRSLIDGKTRATDVMIAGKAVVAGYGVGKCAQALRGFCARVITTEIDPIN 248
Db 250 FDTFTYCDRLSDLSKRTTDMFGGKQVVICGVDGKGCAGSLKGGQGCIVVTEVDPI 309
Qy 249 ALQAAMEGYEVTTMDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHDFEIDV 308

Db 310 ALQAAADGFRVRLNEVIRTVDDVVVATATGNVITRDHNRMKNGCILCNMGHSCSEIDV 369
QY 309 KWLNEAVEKNVLPQVDRYRLKNGRRIILAEGRVLNLCGAMGHPSPFVMSNSFTNQVMA 368
Db 370 NGLHTPELTWVRVRSQVDHHRWPDGRMIILAEGRVLNLSGST-ISSFVSVASSTQALA 428
QY 369 QIELWTHPD-KYPVGVHFLPKKLDDEAVAEHLGKLVNKLTKLTKQAQYLGMSCDGPPKPD 428
Db 429 LIELFAPGRYKSDVLLPKQWDEYVASLHLATFDHAHLTDELTDQSKFMGLNKAGPFKAN 488
QY 429 HVR 431
Db 489 YVR 491

RESULT 8

US-08-896-005-1
; Sequence 1, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,005
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0337 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 500 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT04
CLONE: 1519044
US-08-896-005-1

Query Match 50.5%; Score 1146; DB 2; Length 500;
Best Local Similarity 51.3%; Pred. No. 6,7e-120;
Matches 218; Conservative 84; Mismatches 121; Indels 2; Gaps 2;

QY 9 VADIGLAWGRKALDIAENEMPGLMWRERYASKPLKGARIAGCLHMTVETAVLIETLV 68
Db 77 VKNIKQAEFRRIETAEQMSALISLRKAAQEKPLAGAKIVGCTHTHTQAVLIETLC 136
QY 69 TLGAEVQWSSCNIFSTONHAAAAIAKAGIPVYAWKGETDBEYLWCIEQTLFYFKDGLNMI 128

Db 137 ALGAQCRWSACNIYSTQNEVAAALAEAGVAVPAWKGESDDFMWCIDRCVNMHGWQANMI 196
QY 129 LDGGLDNLHTKYKYPOLLPGIRIGISETTTGVHNLXKMGANGILKVPAINVNDVYTKSK 188
Db 197 LDGGLDTHWCKTIPNVFKIRGIVERSTGVHRLYQLSKAGKLCVPAWNVNDVYTKOK 256
QY 189 FNLVYCCRESLDIGIKRATDVMIAGKAVVAGYGVGVGKCAQALRGFGARVITEIDPIN 248
Db 257 FNLVYCCRESLDGLKRTTDMVFGKQVWVCGYGEVKGCCAAKALGAIVVITEIDPIC 316
QY 249 ALQAAAMEGYVTMBEACQEGNIFVTTCIDIIILGRHPEOMKDDAIVCNIGHFVDEIDV 308
Db 317 ALQACMDGFRVVKLNEVIRQVDVITCTGNKVNVTREHLDRMKNSCIVCNMGHNSIDEV 376
QY 309 KWLNEAVEKNVLPQVDRYRLKNGRRIILAEGRVLNLCGAMGHPSPFVMSNSFTNQVMA 368
Db 377 TSLRTPELTWVRVRSQVDHVRWPDGRVLLAEGRVLNLSGST-VPTFVLSITATTQALA 435
QY 369 QIELWTHPD-KYPVGVHFLPKKLDDEAVAEHLGKLVNKLTKLTKQAQYLGMSCDGPPKPD 427
Db 436 LIELYNAPEGRYKQDVYLLPKQWDEYVASLHLPSFDAHLTELTDQAKYLGMLNKAGPPKP 495
QY 428 DHYR 432
Db 496 NYRY 500

RESULT 9

US-08-204-740-7
; Sequence 7, Application US/08204740
; Patent No. 5753432
; GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,740
FILING DATE: 04-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5753432nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-204-740-7

Query Match 18.5%; Score 421; DB 1; Length 97;

```
Best Local Similarity 100.0%; Pred. No. 1.7e-39;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASAKPLKGARIAGCLHMTVET 60
Db 16 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASAKPLKGARIAGCLHMTVET 75
QY 61 AVLIETLVTLGAEVQWSSCNIF 82
Db 76 AVLIETLVTLGAEVQWSSCNIF 97

RESULT 10
US-09-081-167A-7
; Sequence 7, Application US/09081167A
; Patent No. 6083745
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,167A
; FILING DATE: 18-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083745nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-KK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-081-167A-7

Query Match 18.5%; Score 421; DB 3; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.7e-39;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASAKPLKGARIAGCLHMTVET 60
Db 16 MSDKLPYKVDIGLAAGWKALDIAENEMPGMLMRERYASAKPLKGARIAGCLHMTVET 75
QY 61 AVLIETLVTLGAEVQWSSCNIF 82
Db 76 AVLIETLVTLGAEVQWSSCNIF 97

RESULT 12
US-09-416-833-7
; Sequence 7, Application US/09416833
; Patent No. 6197521
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allgretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
```

; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/416,833
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/204,740
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6197521nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-416-833-7

Query Match 18.5%; Score 421; DB 3; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.7e-39;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDKLPYKVADIGLAAGWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVET 60
DB 16 MSDKLPYKVADIGLAAGWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVET 75
QY 61 AVLIETLVTLGAQVQWSSCNIF 82
DB 76 AVLIETLVTLGAQVQWSSCNIF 97

RESULT 13
PCT-US95-02521-7
; Sequence 7, Application PC/TUS9502521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02521
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-02521-7

Query Match 18.5%; Score 421; DB 5; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.7e-39;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSDKLPYKVADIGLAAGWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVET 60

DB 16 MSDKLPYKVADIGLAAGWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVET 75
QY 61 AVLIETLVTLGAQVQWSSCNIF 82
DB 76 AVLIETLVTLGAQVQWSSCNIF 97
RESULT 14
US-08-204-740-9
; Sequence 9, Application US/08204740
; Patent No. 5753432
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allgretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,740
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5753432nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-204-740-9

Query Match 17.1%; Score 388; DB 1; Length 95;
Best Local Similarity 92.7%; Pred. No. 8.7e-36;
Matches 76; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSDKLPYKVADIGLAAGWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVET 60
DB 14 MSDKLPYKVADIGLAAGWKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVET 73
QY 61 AVLIETLVTLGAQVQWSSCNIF 82
DB 74 AVLIETLVTLGAQVQWSSCNIF 95

RESULT 15
US-09-081-167A-9
; Sequence 9, Application US/09081167A
; Patent No. 6083745
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander

```

; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,167A
; FILING DATE: 18-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083745nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-KK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-081-167A-9

```

```

Query Match      17.1%; Score 388; DB 3; Length 95;
Best Local Similarity 92.7%; Pred. No. 8.7e-36;
Matches 76; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      1 MSDKLPYKVADIGLAANGRKALDIAENEMPGLMREMYASAKPLKGARIAGCLHMTVET 60
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      14 MSDKLPYKVADIGLAANGRKALDIAENEMPGLMREMYASAKPLKGARIAGCLHMTVET 73

Qy      61 AVLIETLVTLGAEVQWSSCNIF 82
        |||||  |||||  |||||  |||||
Db      74 AVLIETKVALGAEARWSSCNIF 95

```

Search completed: December 30, 2003, 10:43:06
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2003, 10:38:13 ; Search time 41 Seconds
(without alignments)
2718.993 Million cell updates/sec

Title: US-10-043-787-1
Perfect score: 2271
Sequence: 1 MSKLPYKVDIGLAAGRK.....QAQLGMSCDGPFKPDHYRY 432

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2190	96.4	432	Q8HXL1	Q8hxl1 sus scrofa
2	1819.5	80.1	432	Q9VXV5	Q9v xv5 drosophila
3	1805.5	79.5	432	Q8MZ11	Q8mz11 drosophila
4	1665	73.3	324	Q8BP17	Q8bp17 mus musculus
5	1522.5	67.0	436	Q8XV1	Q8xv1 leptospira
6	1516	66.8	471	Q8KEG8	Q8keg8 chlorobium
7	1476	65.0	285	Q9H4U6	Q9h4u6 homo sapien
8	1401	61.7	480	Q8PCH5	Q8pch5 xanthomonas
9	1393	61.3	469	Q8GGL7	Q8gg17 streptomyces
10	1386	61.0	480	Q8PP84	Q8pp84 xanthomonas
11	1309.5	57.7	463	Q8KV9	Q8kv9 uncultured
12	1305	57.5	466	Q8FXZ7	Q8fxz7 brucella su
13	1291	56.8	485	Q8LE20	Q8le20 arabidopsis
14	1277	56.2	485	Q949Z9	Q949z9 arabidopsis
15	1274	56.1	485	Q8LPS8	Q8lps8 arabidopsis
16	1274	56.1	485	Q944K5	Q944k5 arabidopsis

17	1274	56.1	485	10	Q9LK36	Q9lk36 arabidopsis
18	1246	54.9	450	10	Q42939	Q42939 nicotiana s
19	1238	54.5	474	16	Q8NSC4	Q8nec4 corynebacte
20	1233	54.3	478	16	Q8FRJ4	Q8frj4 corynebacte
21	1218	53.6	479	5	Q8MUG1	Q8mug1 plasmodium
22	1156.5	50.9	492	5	Q8MYX7	Q8myx7 drosophila
23	1151	50.7	530	4	Q96PK4	Q96pk4 homo sapien
24	1151	50.7	597	4	Q9UG84	Q9ug84 homo sapien
25	1149.5	50.6	521	5	Q9VZX9	Q9vzx9 drosophila
26	1147	50.5	508	11	Q8BIH1	Q8bih1 mus musculu
27	916.5	40.4	431	10	Q947H3	Q947h3 petunia hyb
28	797.5	35.1	425	2	Q8GDW5	Q8gdw5 heliobacill
29	797	35.1	429	16	Q8DGC8	Q8dgc8 synechococc
30	789	34.7	312	10	Q9SDP1	Q9sdp1 allium cepa
31	786	34.6	411	17	Q8PUO4	Q8puq4 methanosaar
32	779	34.3	411	17	Q8TRA5	Q8tra5 methanosaar
33	752	33.1	143	4	Q8BZ13	Q8bz13 homo sapien
34	581	25.6	195	10	Q9X818	Q9xe18 alexandrium
35	527	23.2	181	2	Q9R6R6	Q9r6r6 mycobacteri
36	511	22.5	218	4	Q43210	Q43210 homo sapien
37	450	19.8	202	4	Q9BTL0	Q9bt10 homo sapien
38	449	19.8	232	10	Q8W530	Q8w530 zea mays (m
39	423	18.6	171	10	Q8SP98	Q8sp98 solanum cha
40	423	18.6	500	16	Q8GSA1	Q8gsa1 bifidobacte
41	409.5	18.0	140	2	Q9R6R7	Q9r6r7 mycobacteri
42	377	16.6	121	10	Q41974	Q41974 arabidopsis
43	362	15.9	90	10	Q9M4V0	Q9m4v0 lupinus lut
44	336	14.8	102	10	Q9XF45	Q9xf45 gossypium h
45	324	14.3	96	2	Q93CC0	Q93cc0 mycobacteri

ALIGNMENTS

RESULT 1

Q8HXL1	PRELIMINARY;	PRT;	432 AA.
ID	Q8HXL1		
AC	Q8HXL1;		
DT	01-MAR-2003 (Tremblrel. 23, Created)		
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)		
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)		
DE	S-adenosylhomocysteine hydrolase (EC 3.3.1.1).		
GN	AHCY.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22326073; PubMed=12438749;		
RA	Leeb T.; Rohrer G.A.;		
RT	"Characterization and chromosomal assignment of the porcine AHCY gene		
RT	for S-adenosylhomocysteine hydrolase.";		
RL	Cytogenet. Genome Res. 97:116-119(2002).		
DR	EMBL; AJ422131; CAD19504.1; -.		
KW	Hydrolase.		
SQ	SEQUENCE 432 AA; 47694 MW; 0B756B5F83F6B90E CRC64;		

Query Match	96.4%;	Score 2190;	DB 6;	Length 432;
Best Local Similarity	96.3%;	Pred. NO. 1.4e-159;		
Matches 416;	Conservative	8;	Mismatches	0; Gaps 0;
Indels				
QY	1	MSDKLPYKVDIGLAAGRKALDIENEMPGMLMRERYSASKPLKGARIAGCLHMTVET	60	
Db	1	MSKLPYKVDIGLAAGRKALDIENEMPGMLMRERYSASKPLKGARIAGCLHMTVET	60	
QY	61	AVLIETLVLTGAEVOWSSCNIFSTQHAATAAKAGIPYAWKGETDEBYLWCIEQTLYP	120	
Db	61	AVLIETLVLTGAEVOWSSCNIFSTQHAATAAKAGIPYAWKGETDEBYLWCIEQTLYP	120	
QY	121	KDGPLNMLDDGGDLTNLHTKYPQLLPQIRIGISEBTGTVNLYKWMANGILKVPAINV	180	
Db	121	KDGPLNMLDDGGDLTNLHTKYPQLLPQIRIGISEBTGTVNLYKWMANGILKVPAINV	180	

QY 181 NDSVTKSPDNLYGCRSLIDGKIKRATDVMIAGKAVVAGYGVGVGKCAQALGFGARVI 240
 DB 181 NDSVTKSPDNLYGCRSLIDGKIKRATDVMIAGKAVVAGYGVGVGKCAQALGFGARVI 240
 QY 241 ITEIDPINALQAAMEGYEVTTMDEACQEGNIFVTTTCGIDIIILGRHFEQMKDDAIVCNIG 300
 DB 241 ITEIDPINALQAAMEGYEVTTMDEACQEGNIFVTTTCGIDIIILGRHFEQMKDDAIVCNIG 300
 QY 301 HFVEIDVKNLWNAVEKVNIPQVDYRLKNGRRILLAEGLVNLGCMGHPSPFVMSN 360
 DB 301 HFVEIDVKNLWNAVEKVNIPQVDYRLKNGRRILLAEGLVNLGCMGHPSPFVMSN 360
 QY 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKKLDEAVAEHLGKLVNKLTKTEKQAQYLGM 420
 DB 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKKLDEAVAEHLGKLVNKLTKTEKQAQYLGM 420
 QY 421 CDGPFKPDHRY 432
 DB 421 REGPFKPDHRY 432

RESULT 2
 Q9VXV5 PRELIMINARY; PRT; 432 AA.
 AC Q9VXV5
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Ahcyl13 protein (EC 3.3.1.1) (Adenosylhomocysteinase) (S-adenosyl-L-homocysteine hydrolase) (AdoHcyase).
 GN AHCV13 OR CG11654.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.D.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laekko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-
 CC ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
 CC THEREFORE ADENOSYLHOMOCYSTEINE MAY PLAY A KEY ROLE IN THE
 CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
 CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE
 CC + L-HOMOCYSTEINE.
 CC -!- COPACTOR: NAD (BY SIMILARITY).
 CC -!- PATHWAY: ACTIVATED METHYL CYCLE.
 CC -!- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
 DR EMBL: AE003499; AAF48453.1; --
 DR HSSP: P10760; 1B3R.
 DR FlyBase: FBgn0014455; Abcy13.
 DR InterPro: IPR000043; Ado_hcyase.
 DR Pfam: PF00670; AdoHcyase; 1.
 DR TIGRFAMs: TIGR00936; ahcy; 1.
 DR PROSITE: PS00738; ADOHCYASE_1; 1.
 DR PROSITE: PS00739; ADOHCYASE_2; 1.
 KW Hydrolase; NAD; One-carbon metabolism.
 SQ SEQUENCE 432 AA; 47366 MW; 2977DAF12B40C324 CRC64;

Query Match 80.1%; Score 1819.5; DB 5; Length 432;
 Best Local Similarity 80.7%; Pred. No. 3.5e-131;
 Matches 347; Conservative 28; Mismatches 54; Indels 1; Gaps 1;

QY 4 KLPYKVADIGLAAGKALDIAENEMPGIMRREYSASKPLKGARIAGLHMTVETAVL 63
 DB 3 KESYKVADISLAEGWGRKAIITAEENEMPGIMRREYSASKPLKGARIAGLHMTVETAVL 62
 QY 64 IETLVTLGAEVOWSSCNIFSTONHAAAIATAKAGIPVYAWKGTDEEYLMWICIEQTLVFKDG 123
 DB 63 IETLVTLGAEVOWSSCNIFSTONHAAAIATAKAGIPVYAWKGTDEEYLMWICIEQTLVFKDG 122
 QY 124 -PLNMILDDGGDLTNLHNTKYPQLPGIRIGISETTTGVHNLVYKMWANGILKVPAINVD 182
 DB 123 -PLNMILDDGGDLTNLHNTKYPQLPGIRIGISETTTGVHNLVYKMWANGILKVPAINVD 182
 QY 183 SVTKSKFNLVYKMWANGILKVPAINVD 242
 DB 183 SVTKSKFNLVYKMWANGILKVPAINVD 242
 QY 243 EIDPINALQAAMEGYEVTTMDEACQEGNIFVTTTCGIDIIILGRHFEQMKDDAIVCNIGHF 302
 DB 243 EIDPINALQAAMEGYEVTTMDEACQEGNIFVTTTCGIDIIILGRHFEQMKDDAIVCNIGHF 302
 QY 303 DVEIDVKNLWNAVEKVNIPQVDYRLKNGRRILLAEGLVNLGCMGHPSPFVMSNSF 362
 DB 303 DVEIDVKNLWNAVEKVNIPQVDYRLKNGRRILLAEGLVNLGCMGHPSPFVMSNSF 362
 QY 363 TNOVMAQIELWTHDPKYPVGVHFLPKKLDEAVAEHLGKLVNKLTKTEKQAQYLGMSCD 422
 DB 363 TNOVMAQIELWTHDPKYPVGVHFLPKKLDEAVAEHLGKLVNKLTKTEKQAQYLGMSCD 422
 QY 423 GPFKPDHRY 432
 DB 423 GPFKPDHRY 432

RESULT 3
 Q9VXV5 PRELIMINARY; PRT; 432 AA.
 AC Q9VXV5
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE GM02466p.
 GN AHCV13 OR CG11654.

OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceinker S.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AX102668; AAM27497.1; -.
 DR FlyBase; FBgn0014455; Ahcy13.
 DR InterPro; IPR000043; Ado_hcyase.
 DR Pfam; PF00670; Adohcyase_1.
 DR TIGRfam; TIGR00936; ahcy; 1.
 DR PROSITE; PS00738; ADHCHYASE_1; 1.
 DR PROSITE; PS00739; ADHCHYASE_2; 1.
 SQ SEQUENCE 432 AA; 47338 MW; FEAGB524D88FEB86 CRC64;
 Query Match 79.5%; Score 1805.5; DB 5; Length 432;
 Best Local Similarity 80.2%; Pred. No. 4.1e-130;
 Matches 345; Conservative 29; Mismatches 55; Indels 1; Gaps 1;
 QY 4 KLPYKADIGLAAGKALDAIENEMPLMRERYASKPLKAGARTAGCLHMTVETAVL 63
 DB 3 KPSYKADISLAAGWKALIIIAENEMPLMACRKYGPSKPLKAGARTAGCLHMTVETAVL 62
 QY 64 IETLVTLGAEVQWSSCNIFSTQNHAAIAKAGIPVYANKGETDEEYLWCIEQTLFKDQ 123
 DB 63 IETLVELGAQVQWSSCNIFSTQNHAAIAATGPVYANKGETDEEYMCIEQTLVFPDQ 122
 QY 124 -PLNMILDDGGDLTNLIHTKYPOLLPKIRGISETTTGVNLYKMGANGILKVPAINVND 182
 DB 123 QPLNMILDDGGDLTNLVHKEFPQVLYKNIGLSETTTGVNLYKMPKEGRGLVPAINVND 182
 QY 183 SVTKSFNDLYGCRSLIDGKIKRATDVMIAKGVAVAGVGVGKCAQALRGFGARVIT 242
 DB 183 SVTKSFNDLYGCRSLIDGKIKRATDVMIAKGVAVAGVGVGKCAQALRGFGARVIT 242
 QY 243 EIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGGCIDIILGRHFEQMKDDAIVCNIGHF 302
 DB 243 EVDPIALQAAMEGYEVTMTDEEASKEASIFVTTTGCRTDIITSVHLQMPDDAIVCNIGHF 302
 QY 303 DVEIDVKLNENAVEKNIKPQVDYRLKNGRIILLAEGRVLNLCMGHGPSFVMSNSF 362
 DB 303 DIEIDVDLNLANAKENKVNKPVIRHTMQSGKHILLAEGRVLNLCMGHGPSFVMSNSF 362
 QY 363 TNQVMAQIELWTHPDKYPVGVHFLPKKLDVAEVAHLKLNVLTKLTKEQAOYLGMSCD 422
 DB 363 TNQVLAQIELWTSKDYAVGVHVLPLKLDLBEVASLHLEKLGVLTKLTKEQATYLVGSQT 422
 QY 423 GPFKPDHYRY 432
 DB 423 GPFKPDHYRY 432
 RESULT 4
 Q8BP17 PRELIMINARY; PRT; 324 AA.
 ID Q8BP17
 AC Q8BP17
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE S-adenosylhomocysteine hydrolase (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Body;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The PANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RL Nature 420:563-573 (2002).
 DR EMBL; AK075629; BAC35867.1; -.
 FT NON TER 1
 SQ SEQUENCE 324 AA; 36040 MW; 370014DE8B2EEFDB CRC64;
 Query Match 73.3%; Score 1665; DB 11; Length 324;
 Best Local Similarity 97.2%; Pred. No. 1.1e-119;
 Matches 315; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 109 EYLMCIETQTLFKDGLPLNMILDDGGDLTNLIHTKYPOLLPKIRGISETTTGVNLYKMG 168
 DB 1 EYLMCIETQTLFKDGLPLNMILDDGGDLTNLIHTKYPOLLPKIRGISETTTGVNLYKMG 60
 QY 169 ANGILKVPAINVNDSVTKSFNDLYGCRSLIDGKIKRATDVMIAKGVAVAGVGVGKGC 228
 DB 61 SNGILKVPAINVNDSVTKSFNDLYGCRSLIDGKIKRATDVMIAKGVAVAGVGVGKGC 120
 QY 229 AQUALRGFGARVITTEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIILGRHFE 288
 DB 121 AQUALRGFGARVITTEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIILGRHFE 180
 QY 289 QMKDDAIVCNIGHFDVEIDVKLNENAVEKNIKPQVDYRLKNGRIILLAEGRVLNLC 348
 DB 181 QMKDDAIVCNIGHFDVEIDVKLNENAVEKNIKPQVDYRLKNGRIILLAEGRVLNLC 240
 QY 349 CAMGHPSFVMSNSFTNQVMAQIELWTHPDKYPVGVHFLPKKLDVAEVAHLKLNVLTK 408
 DB 241 CAMGHPSFVMSNSFTNQVMAQIELWTHPDKYPVGVHFLPKKLDVAEVAHLKLNVLTK 300
 QY 409 LTKAQYALGMSCDGPFPKPDHYRY 432
 DB 301 LTKAQYALGMSPINGPFPKPDHYRY 324
 RESULT 5
 Q8EXV1 PRELIMINARY; PRT; 436 AA.
 ID Q8EXV1
 AC Q8EXV1
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE S-adenosylhomocysteine hydrolase.
 GN LB106.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE011599; AANS1665.1; -.
 KW Hydrolyase; Complete proteome.
 SQ SEQUENCE 436 AA; 48233 MW; 1E73AE26ECBDFAC1 CRC64;
 Query Match 67.0%; Score 1522.5; DB 16; Length 436;
 Best Local Similarity 69.2%; Pred. No. 2.1e-108;
 Matches 297; Conservative 48; Mismatches 81; Indels 3; Gaps 2;
 QY 5 LPYKVADIGLAAGKALDAIENEMPLMRERYASKPLKAGARTAGCLHMTVETAVL 64
 DB 10 LSKYKVDLSQAEWGRQEIIAEKEMPLMALAQEYKGGKPLAGARTAGSLHMTVETAVL 69
 QY 65 ETLVTLGAEVQWSSCNIFSTQNHAAIAKAGIPVYANKGETDEEYLWCIEQTLVFPDQ 124
 DB 70 ETLTLGAEVQWSSCNIFSTQNHAAIAKAGIPVYANKGETDEEYLWCIEQTLVFPDQ 129

```

125 LNMILDDGGDLTLNLIHTKYPOLLPGRIGISEETTGTGHNLYKMMANGILKVPAINVNDV 184
130 PNMILDDGGDLTAIHEKYPLLSINGISEETTGTGKSLYKULKGELKVPAINVNDV 189
185 TSKSFNLYGCRSLDGIKRAITDVMIAAGVAVVAGVDVGKCAQALRGFGARVIITEI 244
190 TSKSFNLYGCRSLDGIKRAITDVMIAAGVAVVAGVDVGKCAQALRGFGARVIITEI 249
245 DPINALQAAMEGVEVTMDACOGNIFVTTTCIDIIILGRHPEQMDDAIVCNIGHFDV 304
250 DPICALQASMEGQVLRVEDIIBQVDIVVTATGNDDBIITLEHMKAMKOGAILCNIGHFDT 309
305 EIDVKML-NENAVEKVNIPQVDYRLKNGRRRIILAEGLRNLNLCAMGHPSFVMSNST 363
310 EIQMSRLNEKGVTKKEIKQVDKYTFPDGKSIIVLAEGRLNLCATGHPSPVMSCSFT 369
364 NOVMAQIELWTHDPKYPVGVHFLPKLDEAVAEHLGKLVNLTKEQAQYLGMSCDG 423
370 NOVLAQIELYN--NKVELGVYTLPKHLDKVAALHLEQLGVLRTKLNQADYLVGPING 427
424 PFKPDHYRY 432
428 PFKPDHYRY 436

RESULT 6
Q8KEG8 PRELIMINARY; PRT; 471 AA.
ID Q8KEG8
AC Q8KEG8
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Adenosylhomocysteinase.
GN SAHH OR CT0721.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Winn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.P., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514 (2002).
DR EMBL; AB012843; AAM71958.1; -.
DR TIGR; CT0721; -.
DR InterPro; IPR000043; Ado_hcyase.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF00670; AdoHcyase; 1.
DR TIGRFAM; TIGR00936; ahcy; 1.
DR PROSITE; PS00738; AdoHcyase_1; 1.
DR PROSITE; PS00739; AdoHcyase_2; 1.
KW Complete proteome.
SQ SEQUENCE 471 AA; 51949 MW; 31B082405D241047 CRC64;

```

```

Query Match 66.8%; Score 1516; DB 16; Length 471;
Best Local Similarity 64.2%; Pred. No. 7.4e-108;
Matches 300; Conservative 49; Mismatches 76; Indels 42; Gaps 6;

QY 5 LPYKADIGLAAGWKALDIAENEMPGMLMRERYASKPLKGIAGTGLHMTVETAVLI 64
DB 8 LDYKVADISLAEMGRKEIIEAEKEMPGMLMATRKYEKKPLAGTAGSLHMTIQTAVLI 67
QY 65 ETLVTLGAEVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIEQTLVFKD-- 122

```

```

68 ETLVELGADVRAWSCNIFSTQDHAAAAIAAAGVPVFAWKGETLDEYVMCTROLLEPEGGL 127
123 GPLNMILDDGGDLTLNLIHTKY-----PQLL----- 147
128 GP-NLIVDDGGDTLMIHFGYKIENDPDMLDKTPGNAEEKALLQQLKAVFAEDNQRWKVV 186
148 -PGIGISEETTGTGHNLYKMMANGILKVPAINVNDVTSKSFNLYGCRSLDGIKRA 206
187 AAGMGVSEETTGTGHRLLYQMKGELLPPAIVNVDVTSKSFNLYGCRSLDGIKRA 246
207 TDVMIAGVAVVAGVDVGKCAQALRGFGARVIITEIDPINALQAAMEGVEVTMDAC 266
247 TDVMIAGVAVVAGVDVGKCAQALRGFGARVIITEIDPICALQAAMEGVEVTMDAC 306
267 QGNIPTVTTTCIDIIILGRHPEQMDDAIVCNIGHFDVDEIDVKWLINE-NAVEKVNIPQV 325
307 KEGNIFVTATGNDDBIITLHMKAMKOGAILCNIGHFDVDEIDVKWLINE-NAVEKVNIPQV 366
326 DRYLKNGBRIITLAEGLRNLNLCAMGHPSFVMSNSTNOVMAQIELWTHDPKYPVGVH 385
367 DKYVFENGNCIYLLAEGLRNLNLCATGHPSPVMSNSTNOVMAQIELWTHDPKYPVGVH 424
385 LPKLDDEAVAEHLGKLVNLTKEQAQYLGMSCDGPFKPDHYRY 432
425 LPKLDDEAVAEHLGKLVNLTKEQAQYLGMSCDGPFKPDHYRY 471

```

```

RESULT 7
Q9H4U6 PRELIMINARY; PRT; 285 AA.
ID Q9H4U6
AC Q9H4U6
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)
DE BK3216D2.1.1 (EC 3.3.1.1) (S-adenosylhomocysteinase hydrolase (SAHH),
DE isoform 1) (Adenosylhomocysteinase) (S-adenosyl-L-homocysteinase
DE hydrolase) (AdoHcyase).
GN AHY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Brown A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-
CC ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE
CC + L-HOMOCYSTEINE.
CC -!- COFACTOR: NAD (BY SIMILARITY).
CC -!- PATHWAY: ACTIVATED METHYL CYCLE.
CC -!- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
DR EMBL; AL356299; CAC09529.1; -.
DR HSSP; P10760; 1B3R.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF00670; AdoHcyase; 1.
DR PROSITE; PS00738; AdoHcyase_1; 1.
DR PROSITE; PS00739; AdoHcyase_2; 1.
KW Hydrolase; NAD; One-carbon metabolism.
SQ SEQUENCE 285 AA; 30871 MW; 762B417FAB54FEB3 CRC64;

```

```

Query Match 65.0%; Score 1476; DB 4; Length 285;
Best Local Similarity 99.6%; Pred. No. 4.2e-105;
Matches 284; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDKLPYKVADIGLAAGWKALDIAENEMPGMLMRERYASKPLKGIAGTGLHMTVET 60
DB 1 MSDKLPYKVADIGLAAGWKALDIAENEMPGMLMRERYASKPLKGIAGTGLHMTVET 60

```

```

QY 61 AVLIETLVTLGAQVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEBYLWCIEQTLVF 120
DB 61 AVLIETLVTLGAQVQWSSCNIFSTQDHAAAAIAKAGIPVYAWKGETDEBYLWCIEQTLVF 120

QY 121 KDGPLNMLDDGGDLNLHITKYQPLLPGIRGISEETTTGVHNLKYMANGILKVPAINV 180
DB 121 KDGPLNMLDDGGDLNLHITKYQPLLPGIRGISEETTTGVHNLKYMANGILKVPAINV 180

QY 181 NDSVTKSKFNLKYGCSRESLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
DB 181 NDSVTKSKFNLKYGCSRESLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240

QY 241 ITEIDPINALQAAMEGYEVTTMDAEACQEGNIFVTTTGCIDIIILGR 285
DB 241 ITEIDPINALQAAMEGYEVTTMDAEACQEGNIFVTTTGCIDIIILGR 285

RESULT 8
Q8PCH5 PRELIMINARY; PRT; 480 AA..
AC Q8PCH5;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DE Adenosylhomocysteinease.
GN SAHH OR XCC0752.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MDLINE=22022145; PubMed=12024217;
RA da Silva A.C., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., de Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locail E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AE012174; ANM40067.1; -.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF00670; AdoHcyase; 1.
DR TIGRPFAM; TIGR00936; ahcy; 1.
DR PROSITE; PS00738; ADHCYASE_1; 1.
DR PROSITE; PS00739; ADHCYASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 480 AA; 52629 MW; 2F2B6C776D9B9E54 CRC64;

Query Match 61.7%; Score 1401; DB 16; Length 480;
Best Local Similarity 60.9%; Pred. No. 5e-99;
Matches 285; Conservative 45; Mismatches 96; Indels 42; Gaps 6;

QY 7 YKVADIGLAAGRWKALDIAENMPGLMRERYASAKPLKGARIAGCLHMTVETAVLIET 66
DB 13 YKIDISLADGKRELDAIEHMPGLMSIRKHAQTPLKDVRLTGLSHMTIQTAVLIET 72

QY 67 LVTLGAQVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEBYLWCIEQTLVF--KGGP 124
DB 73 LKDIGANVRWASCNIFSTQDHAAAAIAATGTPVPAWKGETLEEYWDCTLDALTFTLPDGT 132

```

```

QY 125 L---NMILDGGDLNLHITKY-----POLL-----PCI----- 150
DB 133 LTGPVLVDDGGDVTLLIHKGVELENGSTWDEPASSHEEGVIKALLKRVAVVERPGYWAR 192

QY 151 -----RGISEETTTGVHNLKYMANGILKVPAINVNDSTVTKSKFNLKYGCSRESLIDGIIKR 205
DB 193 VVKDWKGVSEETTTGVHRLYQIAEAGKLLIPAINVNDSTVTKSKFNLKYGCSRESLADGLKR 252

QY 206 ATDVMIAGKAVVAGYGVGKCAQALRGFGARVIITEIDPINALQAAMEGYEVTTMDAE 265
DB 253 ADMVMLAGKAVVCGYGVGKGSASLRAIGARVIVTEIDPICALQASMEGFEVNTIEST 312

QY 266 COEGNIFVTTTGCIDIIILGRHFEOMKDDAIVCNIGHFDVDEIDVKWLN--NAVYKVNIKPQ 324
DB 313 LGRGDIYVTTTGNKDIITVEHLQAMKDKQAIVCNIGHFDNEIQVDALNALKVEKINIKPQ 372

QY 325 VDRYRLONGRRRIILAEGRNLVNLGCMGHPSPVMSNPTNQVMAQIBLTHTPDKYPVGVI 384
DB 373 VDKYTVFGNGNAIFLLADGRLNLGCATGHPSPVMSNFANQTLAQIDLWKRDSYEKKVY 432

QY 385 FLPKKLDEAAEAHGLKNVLTKLTEKQAQYLGMSCDGPPKPDHRY 432
DB 433 ILPKHLDEEVARLHLEKIGVKUTLTLDQADYLGVDVAGPYKPDHRY 480

RESULT 9
Q8GGL7 PRELIMINARY; PRT; 469 AA.
AC Q8GGL7;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Adenosylhomocysteinease.
OS Streptomyces atroolivaceus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=66869;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22336326; PubMed=12446651;
RA Cheng Y.Q., Tang G.L., Shen B.;
RT "Identification and Localization of the Gene Cluster Encoding
RT Biosynthesis of the Antitumor Macrolactam Leinamycin in Streptomyces
RT atroolivaceus S-140";
RL J. Bacteriol. 184:7013-7024(2002).
DR EMBL; AF484556; AAN85548.1; -.
SQ SEQUENCE 469 AA; 51088 MW; 9B1980B1A5FCFA2E CRC64;

Query Match 61.3%; Score 1393; DB 2; Length 469;
Best Local Similarity 60.8%; Pred. No. 2e-98;
Matches 282; Conservative 58; Mismatches 90; Indels 34; Gaps 6;

QY 2 SDKLPYKVADIGLAAGRWKALDIAENMPGLMRERYASAKPLKGARIAGCLHMTVETA 61
DB 7 ADFTDPKVADLSLAAGFGRKEITLAEHMPGLMSIREEYAAAOPLAGARITGLHMTVQTA 66

QY 62 VLITLVTLGAQVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEBYLWCIEQTLVF- 120
DB 67 VLITLVTLGAQVQWSSCNIFSTQDHAAAAIAAAGIPVPAWKGETLEEYWMCTEQALTWP 126

QY 121 -KDGPLNMLDDGGDLNLHITKYQPLLPGIRGISEETTTGVHNLKYMANGILKVPAINV 149
DB 127 GHTGP-NMILDGGDVTLLIHKGVEYRKTGILPEAENBELAVVRALLDRSGLDWTANSE 185

QY 150 IRGISEETTTGVHNLKYMANGILKVPAINVNDSTVTKSKFNLKYGCSRESLIDGIIKRATDV 209
DB 186 IRGVTEETTTGVHRLYEMHRDGTLLFPAINVNDATVTKSKFDKNKYGCRHSLIDGINRATDV 245

QY 210 MIAGKAVVAGYGVGKCAQALRGFGARVIVTEIDPINALQAAMEGYEVTTMDAEACQEG 269
DB 246 LIGGKTAVVCGYGVGKGSASLRAIGARVIVTEIDPICALQAAADGYSQVATLDEYVDKA 305

```



```
Db 304 IFITGKNDVIRIEHMRDKMAIYVGNIGHFDNEIQVASLKNHKW-----TNKEQ 355
Qy 325 VDRYLKNGRRIILLAEGRVNLVLCGAMGHPSPFVMSNSTNQVMAQIELWTHDPKYPGVGH 384
Db 356 VDMIEPNNGRLILLSEGRLLNLGNATGHPSPFVMSASFTNQVLAQIELWTKGDDYSNEVY 415
Qy 385 FLPKLDEAVAEHLGKLNKLTKEKQAOYLGMSCDGPFPKPDHYRY 432
Db 416 ILPKHLDKVARLHLDRIKGLSKLNDEQAAYIGVSSEGPFKEHYRY 463

RESULT 12
Q8FXZ7
ID Q8FXZ7 PRELIMINARY; PRT; 466 AA.
AC Q8FXZ7;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Adenosylhomocysteinase.
GN AHY OR BR2097.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.B., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Bean M.J.,
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts."
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
DR EMBL; AE014496; AAN30987.1; -.
DR TIGR; BR2097; -.
KW Complete proteome.
SQ SEQUENCE 466 AA; 50791 MW; 6AB3B4B1C2F0B541 CRC64;

Query Match 57.5%; Score 1305; DB 16; Length 466;
Best Local Similarity 57.4%; Pred. No. 1.le-91;
Matches 272; Conservative 51; Mismatches 101; Indels 50; Gaps 6;

Qy 1 MSKLPYKVADIGLAAGRKALDIENEMPLMRERYASKPLKGARIAGCLHMTVET 60
Db 1 MTASQDFVVKDISLADWGRKELOIAETEMPLMAAREEFGKSQLKGARISGLHMTIQT 60
Qy 61 AVLIETLVLTGASVQWSSCHIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLWCIEBTLF 120
Db 61 AVLIETLVLTGASVQWSSCHIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLWCIEBTLF 120
Qy 121 KDG-PLNMILDDGGDLTNLIH-----HTKYPQL-----LPG--- 149
Db 121 PDGEPNMLDDGGDLTNLIHARAEAGEDVLNPSQSEEEVLPQIKRMAATGFFT 180
Qy 150 -----IRGISEETTTGVNHYLYKWMANGILKVPAINVNDSTKSKFDNLYGCRSLDGIK 204
Db 181 KQRAAIKGVTEETTTGVNHYLYKWMANGILKVPAINVNDSTKSKFDNLYGCRSLDGIK 240
Qy 205 RATDVMIAKGVAVAGVGVGKCAQALRGFGARVITIEDIPINALQAAMEGEVVTMD 264
Db 241 RGTDMVMAGKAVVAGVGVGKCAQALRGFGARVITIEDIPINALQAAMEGEVVTMD 300
Qy 265 ACQEGNIFVTTTGCIDIIILGRHFEQKDDAIVCNIGHFDVEIDV-----KLNENAVEK 318
Db 301 AASTADIVTTTGNKVDITTDHMRKDKMCIKVDIKNIGHFDNEIQVAAIRLNK 352
Qy 319 VNIKPOVDRLNKGRIILLAEGRVNLVLCGAMGHPSPFVMSNSTNQVMAQIELWTHDPK 378
Db 319 VNIKPOVDRLNKGRIILLAEGRVNLVLCGAMGHPSPFVMSNSTNQVMAQIELWTHDPK 378
```

```
Db 353 TNVKEPQVDLIEPPDGKRLILLSEGRLLNLGNATGHPSPFVMSASFTNQVLAQIELFTRTDA 412
Qy 379 YPVGVHFLPKLDEAVAEHLGKLNKLTKEKQAOYLGMSCDGPFPKPDHYRY 432
Db 413 YKNEVTVLPKHLDEKVARLHLDKGLAKLTVLSEQAAYIGVTPOQFPKSEHYRY 466

RESULT 13
Q8LE20
ID Q8LE20 PRELIMINARY; PRT; 485 AA.
AC Q8LE20;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Adenosylhomocysteinase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurooids II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.B.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."
RL Genome Biol. 0:0-0 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085669; AAM62888.1; -.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF00670; AdoHcyase; 1.
DR TIGRFAMs; TIGR00936; ahcy; 1.
DR PROSITE; PS00738; ADOHCYASE_1; 1.
DR PROSITE; PS00739; ADOHCYASE_2; 1.
SQ SEQUENCE 485 AA; 53401 MW; 11132BBC1F46C86C CRC64;

Query Match 56.8%; Score 1291; DB 10; Length 485;
Best Local Similarity 56.3%; Pred. No. 1.4e-90;
Matches 267; Conservative 54; Mismatches 103; Indels 50; Gaps 6;

Qy 7 YKVAIDIGLAAGRKALDIENEMPLMRERYASKPLKGARIAGCLHMTVETAVLIET 66
Db 14 YKVKDMSQADFGRLELEAEVEMPGLMACRTEFGSPQPKGARITGSLHMTITQTAVLIET 73
Qy 67 LVTLAGVQWSSCHIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLWCIEBTL-YPKDGPL 125
Db 74 LTALGAEVQWSSCHIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLWCIEBTL-YPKDGPL 133
Qy 126 NMILDGGDLTNLIH-----HTKYPQL-----LPG--- 144
Db 134 DLIVDDGGDLTNLIH-----HTKYPQL-----LPG--- 193
Qy 145 QLLPGIRISEETTTGVNHYLYKWMANGILKVPAINVNDSTKSKFDNLYGCRSLDGIK 204
Db 194 KMKRLVGVSEETTTGVNHYLYKWMANGILKVPAINVNDSTKSKFDNLYGCRSLDGIK 253
Qy 205 RATDVMIAKGVAVAGVGVGKCAQALRGFGARVITIEDIPINALQAAMEGEVVTMD 264
Db 254 RATDVMIAKGVAVAGVGVGKCAQALRGFGARVITIEDIPINALQAAMEGEVVTMD 313
Qy 265 ACQEGNIFVTTTGCIDIIILGRHFEQKDDAIVCNIGHFDVEIDVKNLNE-NAVEKVNIPK 323
Db 314 VVSEADIVTTTGNKVDITTDHMRKDKMCIKVDIKNIGHFDNEIQVAAIRLNK 373
Qy 324 QVDRY-----RLKNGRIILLAEGRVNLVLCGAMGHPSPFVMSNSTNQVMAQIELWTHDPK 378
Db 374 QVDRY-----RLKNGRIILLAEGRVNLVLCGAMGHPSPFVMSNSTNQVMAQIELWTHDPK 378
Qy 374 QVDRY-----RLKNGRIILLAEGRVNLVLCGAMGHPSPFVMSNSTNQVMAQIELWTHDPK 378
Db 374 QVDRY-----RLKNGRIILLAEGRVNLVLCGAMGHPSPFVMSNSTNQVMAQIELWTHDPK 378
```

379 YPVGVHFLPKKLDEAVAHGLKLVNKLTKLTSKQAYLQMSCGDFKPDHYRY 432
| : : ||| ||| ||| ||| : : : : ||| |||
432 YEKVVYLPFKHLDKVALHGLKGLARLTKLSDQXDYVSIPTEGPKPPHYRY 485

RESULT 14

Q94929 PRELIMINARY; PRT; 485 AA.

AC Q94929;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2002 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2001 (TrEMBLrel. 22, Last annotation update)

DE Putative S-adenosyl-L-homocysteine hydrolase (EC 3.3.1.1)

DE (Adenosylhomocysteinease) (S-adenosyl-L-homocysteine hydrolase)

DE (AdoHCyase).

DE MW9.16.

GN Arabidopsis thaliana (Mouse-ear cress).

OS Arabidopsi

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi

OX NCBI_TaxID=3702;

ON [1]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno P., Dale J.M.,

RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,

RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,

RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B.,

RA Lin J., Meyers M.C., Miranda M., Narusea M., Nguyen M., Palm C.J.,

RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,

RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;

RT "Full length cDNA of gene MYM9.16 (GI:9293955).";

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS; THEREFORE ADENOSYLHOMOCYSTEINE MAY PLAY A KEY ROLE IN THE CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE + L-HOMOCYSTEINE.

CC -1- COFACTOR: NAD (BY SIMILARITY).

CC -1- PATHWAY: ACTIVATED METHYL CYCLE.

CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.

DR EMBL; AY050783; AAK92718.1; -.

DR InterPro; IPR000043; Ado_hcyase.

DR Pfam; PF00670; AdoHcyase; 1.

DR TIGRfams; TIGR00936; ahcy; 1.

DR PROSITE; PS00738; ADOHCYASE_1; 1.

DR PROSITE; PS00739; ADOHCYASE_2; 1.

DR Hydrolase; NAD; One-carbon metabolism.

SQ SEQUENCE 485 AA; 53178 MW; 876079F4782F532 CRG64;

Query Match	56.2%	Score 1277;	DB 10;	Length 485;.
Best Local Similarity	55.5%;	Pred. No. 1.6e-89;		
Matches 262;	Conservative 58;	Mismatches 106;	Indels 46;	Gaps 5;
QY	7	YKVADIGLAAWGRKALDIAENEMPLMRMRYSASKPLKGARTAGCLHMTVETAVLIET	66	
DB	14	YKVKMSQADFGRLIELEAEVEMPLVSVCTEFGPSQPLKGARTGSLHMTIQTAVLIET	73	
QY	67	LVTLGARVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEEVLWCIEBOTL-YFKDQGPL	125	
DB	74	LTLGAEVRWCSCNIFSTQDHAAAAIARDAAAFWAKGETLQEYWWCTERALDWMGPGGPP	133	
QY	126	NMILDDGGDLTNLIH-----TKYP	144	
DB	134	DLIVDDGGDATLLIHGCVKAEIIPAKNGTFDPDTSTDNPEFQIVLSIIKDLQVDPKKYH	193	
QY	145	QLLIPGIRGISSEETTTGVHNIYKMWANGILKVPAINVNDVSTKSKFDNLYGCRSLIDGIK	204	
DB	194	KMKERLVGVSEETTTGVKRYLYQMOETGALLFPAINVNDVSTKSKFDNLYGCRSLPDGLM	253	
QY	205	RATDVNMIAGKVAVVAGYGDVKGCAQALRGFGARVITIEDIPINALQAAHEGVEVTTMDE	264	

DB	254	RATDVMIAKGVAVICGYGDVGKGAAMKTAGARVIVTEIDPICALQALMEGLQVLTLED	313
QY	265	ACQEGNIFFVTTGCTIDIIILGRHFEOMKDDAIVCNIGHGFDEVIDVKWINE-NAVEKVNKP	323
DB	314	VVSEADIFECTTGNKDIIIVDMRMRKKNNAIVCNIGHFEDNIDMLGLETPYGVKRTIKP	373
QY	324	QVDRYRLKN-GRRILLAEGLVNLGCMAGHPFSFVMSNSFTNVMAQIELWTHPD--KYP	380
DB	374	QTRDWVPDNTSGIIVLAEGRLLMNLGCATGHPSFVMSCSFTNOVIAQELMNEKSSGKYE	433
QY	381	GVGHFLPKLDEANAEHLGKLNVLTKLTKEQAOYLGNMCDGPKPDHYR	432
DB	434	KKVYVLPKRLDEKVAALHLGKGLARLTCLKDQSDYVIPVEGPKYPVHYR	485
RESULT 15			
QBLP58	ID	Q8LP58	PRELIMINARY; PRT; 485 AA.
AC	Q8LP58;		
DT	01-OCT-2002	(TrEMBLrel. 22, Created)	
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	AT3g23810/MYm9.15		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.		
OX	NCBI_TaxID=3702;		
RN	[1]	SEQUENCE FROM N.A.	
RP	Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,		
RA	Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,		
RA	Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,		
RA	Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,		
RA	Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,		
RA	Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,		
RA	Yanamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,		
RA	Ecker J.R.;		
RT	"Arabidopsis cDNA clones";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY094404; AAM19782.1; -.		
DR	InterPro: IPR000043; Ado_hcyase.		
DR	Pfam; PF00670; AdoHcyase_1.		
DR	TIGRFAMs; TIGR00936; ahcy; 1.		
DR	PROSITE; PS00738; ADOHCYASE 1; 1.		
DR	PROSITE; PS00739; ADOHCYASE 2; 1.		
SO	SEQUENCE 485 AA; 53101 MW; 142615500996059 CRC64;		

Query Match	56.1%;	Score 1274;	DB 10;	Length 485;
Best Local Similarity	55.5%;	Pred. No. 2.8e-89;		
Matches 262; Conservative	57;	Mismatches 107;	Indels 46;	Gaps 5;
Qy	7	YKVADIGLAAWGRKALDIAENEMPGLMRMERYASAKPLKGARIAGCLHMTVETAVLIET	66	
Db	14	YKKWNSQADFRLIELEAEVEVEGLVCVTERGSPQLKGARITGSLSHWIQTAVLIET	73	
Qy	67	LVTLGAEVQWSSCNIFSTQHAAAAIAKAGIPVYAWKGETDBEYLWCIEQTL-YFKDGPL	125	
Db	74	LTALGAERVWCSCNIFSTQDHAAAATARDSAAVFANKGETLQBYWWCTERALDWGPGGP	133	
Qy	126	NMILDDGDGLTNLIH-----TKYP	144	
Db	134	DLIVDDGDGDTLLIHGVKAELFAKNGTFFDPTSDNPESFQIVLSIKDGLQVDPPKYH	193	
Qy	145	QLLPFGIRGESETTTGVHNYKMWANGILKVPAINVNDSVTSKSFNDLYCRESLIIDGIK	204	
Db	194	KMKERLVGVSEETTGVKRVYQMETGALLFPALINVNDSVTSKSFNDLYICRHSLLPDGLM	253	
Qy	205	RATDVMIAGKVAVVAGYGDKGCACQAIRGFGARVIITEIDPNALQAAMHEGYEVITMD	264	
Db	254	RATDVMIAGKVAVICGYGDKGCACAAKMTAGARVITEIDPICALQAMEGLQVLTE	313	

Qy	265	ACOEGRIFVTTTCIDIIILGRHFEQMKDDAIVCNIGHFDVEIDVKWLN-NAVEKYNIXP	323
Db	314	VVSEADIFCTTTGNKDIIIMVDHMRKKNNAIVCNIGHFDNEIDMLGLETPGVKRIKIP	373
Qy	324	QVDRYRLKN-GREIILLAEGRLVNLGCAMGHPSPVMSNSFTNOVMAQIELWTHPD--KYP	380
Db	374	QVDRWVPDNTSGIIVLAEGRLMNLGCATGHPSPVMSCSFTNQVIAQLELWNEKSSGKYE	433
Qy	381	GVVHFLPKLDEAVAEAHILGKLVNKLTKLTEKQAQYLGMSCDGPPKPDHYRY	432
Db	434	KKYVVLPKHLDERKVAALHLGKLGARLTKUTKDQSDYVSIPIVEGPKPVHRY	485

Search completed: December 30, 2003, 10:42:00
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2003, 10:37:43 ; Search time 17 Seconds
(without alignments)
1195.032 Million cell updates/sec

Title: US-10-043-787-1

Perfect score: 2271

Sequence: 1 MSDKLPYKVADIGLAAGRK.....QAQYLGMSCDGPFKPDHYRY 432

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Description
1	2266	99.8	P23526 homo sapien
2	2193	96.6	P50247 mus musculus
3	2194	96.2	P10760 rattus norv
4	2017.5	88.8	P51893 xenopus lae
5	2015.5	88.7	O93477 xenopus lae
6	1797	79.1	O27580 drosophila
7	1767.5	77.8	O76757 anopheles g
8	1748	77.0	P27604 caenorhabdi
9	1733	76.3	O13639 schizosacch
10	1697.5	74.7	P10819 dictyosteli
11	1633	71.9	P39954 saccharomyc
12	1607.5	70.8	P36889 leishmania
13	1536	67.6	Q12663 pneumocysti
14	1516	66.8	O8keg8 chlorobium
15	1426	62.8	O9pej1 xylella fas
16	1401	61.7	O8pch5 xanthomonas
17	1387	61.1	O8Y387 ralestonia s
18	1386	61.0	O8p84 xanthomonas
19	1373	60.5	O9kxmi streptomyce
20	1347.5	59.3	O9366 streptomyce
21	1335.5	58.8	O50562 rhodobacter
22	1323.5	58.3	P21803 rhodobacter
23	1323	58.3	P35007 catbacteri
24	1315.5	57.9	O08364 mycobacteri
25	1308	57.6	Q92tc1 rhizobium m
26	1306	57.5	Q01781 petroselinu
27	1302.5	57.4	O9ccj4 mycobacteri
28	1302	57.3	O8ye49 brucella me
29	1302	57.3	O9ep37 lupinus lut
30	1299	57.2	O8u399 agrobacteri
31	1299	57.2	P50248 nicotiana t
32	1296	57.1	O98cm3 rhizobium l
33	1295	57.0	P50246 medicago sa

RESULT 1
SAHH_HUMAN
ID SAHH_HUMAN STANDARD; PRT; 432 AA.
AC P23526; O96A36;
DT 01-NOV-1991 (Rel. 20, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenosylhomocysteinease (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN AHCY OR SAHH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=90087640; PubMed=2596825;
RA Coulter-Karis D.E., Herzhfeld M.S.;
RT "Sequence of full length.cdna for human S-adenosylhomocysteine
RT hydrolase.";
RL Ann. Hum. Genet. 53:169-175(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekesch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leharplait M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.B., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams B.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta, and Skin;
RX MEDLINE=22388257; PubMed=12477932;

ALIGNMENTS

34	1293	56.9	462	1	SAHH ROSDE	Q9naa5 roseobacter
35	1293	56.9	485	1	SAHH MESCR	P93253 mesembryant
36	1292	56.9	485	1	SAHH ARATH	O23255 arabidopsis
37	1279	56.3	485	1	SAHH PHASS	P50249 phalaenopsi
38	1277	56.2	485	1	SAHH WHEAT	P32112 triticum ae
39	1256.5	55.3	463	1	SAHH CAUCR	Q9abho caulobacter
40	1256	55.3	485	1	SAHH LYCES	P9abw5 lycopersico
41	1251	55.1	486	1	SAHH TRIVA	P51540 trichomonas
42	1249.5	55.0	469	1	SAHH PSEAB	Q91685 pseudomonas
43	1238	54.5	474	1	SAHH CORGL	Q8nac4 corynebacte
44	1215	53.5	479	1	SAHH PLAP7	P50250 plasmodium
45	1156.5	50.9	492	1	SAHH2 DROME	P50245 drosophila

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RC TISSUE=Placenta;
RX MEDLINE=98246411; PubMed=9586999;
RA Turner M.A., Yuan C.S., Borchardt R.T., Hereshfield M.S., Smith G.D.,
RA Howell P.L.;
RT "Structure determination of selenomethionyl S-adenosylhomocysteine
RT hydrolase using data at a single wavelength.";
RL Nat. Struct. Biol. 5:369-376(1998).
CC -I- FUNCTION: ADENOSYLOHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLOHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLOHOMOCYSTEINE.
CC -I- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -I- COFACTOR: NAD.
CC -I- PATHWAY: Activated methyl cycle.
CC -I- SUBUNIT: Homotetramer.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- DISEASE: DEFICIENCY IN AHCY IS ONE THE DIFFERENT CAUSES OF
CC HYPERMETHIONINEMIA.
CC -I- SIMILARITY: Belongs to the adenosylhomocysteine family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M61831; AAA51681.1; -;
DR EMBL; M61832; AAA51682.1; -;
DR EMBL; AL356299; CAC09528.1; -;
DR EMBL; BC010018; AAH10018.1; -;
DR EMBL; BC011606; AAH11606.1; -;
DR PIR; A43629; A43629.
DR PDB; 1A7A; 20-APR-99.
DR Genew; HGNC:343; AHCY.
DR MIM; 180960; -;
DR GO; GO:0005737; Cytoplasm; NAS.
DR GO; GO:0004013; F:adenosylhomocysteine activity; TAS.
DR GO; GO:0006730; P:one-carbon compound metabolism; NAS.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF05221; AdoHcyase; 1.
DR TIGRfam; TIGR00936; ahcy; 1.
DR PROSITE; PS00738; ADOHCYASE.1; 1.
DR PROSITE; PS00739; ADOHCYASE.2; 1.
KW Hydrolase; NAD; One-carbon metabolism; Polymorphism; 3D-structure.
FT NP_BIND 214 245 NAD (POTENTIAL).
FT VARIANT 86 86 D -> N.

/FTID=VAR_006934.

FT STRAND 8
FT HELIX 12
FT HELIX 15
FT TURN 26
FT TURN 28
FT HELIX 30
FT TURN 39
FT TURN 40
FT TURN 42
FT TURN 44
FT STRAND 47
FT STRAND 49
FT HELIX 53
FT TURN 58
FT TURN 69
FT STRAND 71
FT STRAND 73
FT HELIX 77
FT TURN 86
FT TURN 94
FT TURN 95
FT STRAND 96
FT HELIX 99
FT STRAND 100
FT HELIX 107
FT STRAND 118
FT STRAND 119
FT TURN 120
FT TURN 121
FT STRAND 122
FT STRAND 123
FT STRAND 124
FT STRAND 127
FT STRAND 130
FT HELIX 134
FT TURN 144
FT HELIX 146
FT STRAND 147
FT STRAND 152
FT HELIX 155
FT STRAND 158
FT TURN 169
FT TURN 170
FT STRAND 171
FT TURN 177
FT TURN 181
FT HELIX 184
FT TURN 187
FT TURN 188
FT STRAND 189
FT TURN 190
FT TURN 207
FT TURN 212
FT TURN 213
FT STRAND 215
FT HELIX 223
FT TURN 236
FT STRAND 238
FT STRAND 242
FT TURN 256
FT TURN 258
FT STRAND 259
FT HELIX 262
FT TURN 266
FT STRAND 271
FT STRAND 274
FT STRAND 282
FT STRAND 284
FT TURN 287
FT TURN 288
FT TURN 292
FT STRAND 293
FT STRAND 294
FT STRAND 306
FT HELIX 308
FT STRAND 314
FT STRAND 315
FT TURN 322
FT TURN 324
FT STRAND 325
FT TURN 330
FT STRAND 331
FT STRAND 332
FT STRAND 335
FT STRAND 339
FT HELIX 340
FT HELIX 342
FT HELIX 345
FT HELIX 348
FT HELIX 355
FT HELIX 374
FT STRAND 376
FT STRAND 383
FT HELIX 384
FT TURN 388
FT TURN 397
FT TURN 399
FT HELIX 400
FT TURN 402
FT TURN 403
FT HELIX 411
FT TURN 417
FT TURN 418
FT TURN 428
FT TURN 429
SQ SEQUENCE 432 AA; 47716 MW; 2833C025F969553E CRC64;

Query Match 99.8%; Score 2366; DB 1; Length 432;
Best Local Similarity 99.8%; Pred. No. 1.3e-162;
Matches 431; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDKLPYKVADIGLAAGWKALDIAENEMPGLMRERYASAKPLKGARIAGLHMTVET 60

```

Db 1 MSDKLPYKVADIGLAANGKALDIAENEMPGLMRERYSASKPLKGARIAGCLHMTVET 60
Qy 61 AVLIETLVTLGAEVQSSCNIFSTONHAAATAKAGIPVYANKGETDEEYLWCIETLYF 120
Db 61 AVLIETLVTLGAEVQSSCNIFSTONHAAATAKAGIPVYANKGETDEEYLWCIETLYF 120
Qy 121 KDGPLNMLDDGGDLTNLIHTKYPQLLPGRIGISETTTTGVNLYKWMANGILKVPAINV 180
Db 121 KDGPLNMLDDGGDLTNLIHTKYPQLLPGRIGISETTTTGVNLYKWMANGILKVPAINV 180
Qy 181 NDSVTKSKFDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
Db 181 NDSVTKSKFDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVI 240
Qy 241 ITEIDPINALQAAMEGYEVTWDEACQEGNIFVTTTGCIDIIILGRHFEOMKDAIVCNIG 300
Db 241 ITEIDPINALQAAMEGYEVTWDEACQEGNIFVTTTGCIDIIILGRHFEOMKDAIVCNIG 300
Qy 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGLVNLGCAMGHPSPVMSN 360
Db 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGLVNLGCAMGHPSPVMSN 360
Qy 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKKLDEAVAHAHLGKLVNKLTKLTKTEKQAQYLGMS 420
Db 361 SFTNQVMAQIELWTHDPKYPVGVHFLPKKLDEAVAHAHLGKLVNKLTKLTKTEKQAQYLGMS 420
Qy 421 CDGPPKPDHYRY 432
Db 421 CDGPPKPDHYRY 432

RESULT 2
SAHH_MOUSE
ID SAHH_MOUSE STANDARD; PRT; 431 AA.
AC P50247; Q91W1;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenosylhomocysteinease (EC 3.1.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase) (Liver copper binding protein) (CUBP).
GN AHCY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=95386522; PubMed=7657650;
RA Bethin K.E., Petrovic N., Ettinger M.J.;
RT "Identification of a major hepatic copper binding protein as S-
RT adenosylhomocysteine hydrolase.";
RL J. Biol. Chem. 270:20698-20702(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```

```

RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: ADENOSYLOHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC THEREFORE ADENOSYLOHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLOHOMOCYSTEINE.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: Activated methyl cycle.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the adenosylhomocysteinease family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: L32836; AAA70378.1; -.
CC DR EMBL: BC015304; AAH15304.1; -.
CC DR HSSP: E10760; 1D4F.
CC DR MGD: MGI:87968; Ahcy.
CC DR InterPro: IPR000043; Ado_hcyase.
CC DR Pfam: PF05221; AdoHcyase; 1.
CC DR Pfam: PF00670; AdoHcyase_NAD; 1.
CC DR TIGRFAMs: TIGR00936; ahcy; 1.
CC DR PROSITE: PS00738; ADOHCYASE_1; 1.
CC DR PROSITE: PS00739; ADOHCYASE_2; 1.
CC DR Hydrolase; NAD; One-carbon metabolism.
CC FT INIT MET 0 0
CC FT NP BIND 213 244 NAD (POTENTIAL).
CC FT CONFLICT 173 173 K -> N (IN REF. 1).
CC FT SEQUENCE 431 AA; 47557 MW; D577DB2F9D98CA94 CRC64;
CC
CC Query Match 96.6%; Score 2193; DB 1; Length 431;
CC Best Local Similarity 96.8%; Pred. No. 3.8e-157;
CC Matches 417; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
Qy 2 SDKLPYKVADIGLAANGKALDIAENEMPGLMRERYSASKPLKGARIAGCLHMTVETA 61
Db 1 SDKLPYKVADIGLAANGKALDIAENEMPGLMRERYSASKPLKGARIAGCLHMTVETA 60
Qy 62 VLIETLVTLGAEVQSSCNIFSTONHAAATAKAGIPVYANKGETDEEYLWCIETLYPK 121
Db 61 VLIETLVTLGAEVQSSCNIFSTONHAAATAKAGIPVYANKGETDEEYLWCIETLYPK 120
Qy 122 DGPLNMLDDGGDLTNLIHTKYPQLLPGRIGISETTTTGVNLYKWMANGILKVPAINV 181
Db 121 DGPLNMLDDGGDLTNLIHTKYPQLLPGRIGISETTTTGVNLYKWMANGILKVPAINV 180
Qy 182 DSVTKSKFDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVII 241
Db 181 DSVTKSKFDNLGCRSLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVII 240
Qy 242 TEIDPINALQAAMEGYEVTWDEACQEGNIFVTTTGCIDIIILGRHFEOMKDDAIVCNIGH 301
Db 241 TEIDPINALQAAMEGYEVTWDEACQEGNIFVTTTGCIDIIILGRHFEOMKDDAIVCNIGH 300
Qy 302 FVDEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGLVNLGCAMGHPSPVMSNS 361
Db 301 FVDEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGLVNLGCAMGHPSPVMSNS 360
Qy 362 FTNQVMAQIELWTHDPKYPVGVHFLPKKLDEAVAHAHLGKLVNKLTKLTKTEKQAQYLGMS 421
Db 361 FTNQVMAQIELWTHDPKYPVGVHFLPKKLDEAVAHAHLGKLVNKLTKLTKTEKQAQYLGMS 420

```

QY 422 DGPFPDHYRY 432
DB 421 NGFPKPDHYRY 431

RESULT 3
SAHH RAT
ID SAHH RAT STANDARD; PRT; 431 AA.
AC P10760;
AT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
hydrolase) (AdoHcyase).
GN AHCY.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=87118240; PubMed=3027698;
RA Ogawa H., Gomi T., Mueckler M.M., Fujioka M., Backlund P.S. Jr.,
RA Aksanit R.R., Unson C.G., Cantoni G.L.,
RT "Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from rat
liver as derived from the cDNA sequence".
RL Proc. Natl. Acad. Sci. U.S.A. 84:719-723 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=95262723; PubMed=7744082;
RA Merta A., Aksanit R.R., Kasir J., Cantoni G.L.,
RT "The gene and pseudogenes of rat S-adenosyl-L-homocysteine
hydrolase".
RL Eur. J. Biochem. 229:575-582 (1995).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=99315209; PubMed=10387078;
RA Hu Y., Komoto J., Huang Y., Gomi T., Ogawa H., Takata Y., Fujioka M.,
RA Takusagawa F.,
RT "Crystal structure of S-adenosylhomocysteine hydrolase from rat
liver".
RL Biochemistry 38:8323-8333 (1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=20493601; PubMed=10913437;
RA Komoto J., Huang Y., Gomi T., Ogawa H., Takata Y., Fujioka M.,
RA Takusagawa F.,
RT "Effects of site-directed mutagenesis on structure and function of
recombinant rat liver S-adenosylhomocysteine hydrolase. Crystal
structure of D244E mutant enzyme".
RL J. Biol. Chem. 275:32147-32156 (2000).
CC -!- FUNCTION: ADENOSYLMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
THEREFORE ADENOSYLMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CONCENTRATION OF ADENOSYLMOCYSTEINE.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
+ L-homocysteine.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: Activated methyl cycle.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC CC
CC EMBL; M15185; AAA40705.1; -
CC EMBL; U14937; AAA92043.1; -
CC PIR; A26583; A26583.
CC PIR; 1B3R; 23-DEC-98.
CC PDB; 1D4F; 17-JAN-01.
CC PDB; 1D4G; 17-JAN-01.
CC PDB; 1KOU; 17-OCT-01.
CC PDB; 1KY4; 25-SEP-02.
CC PDB; 1KY5; 25-SEP-02.
CC InterPro; IPR000043; Ado hcyase.
CC Pfam; PF05221; AdoHcyase; 1.
CC TIGRFAMs; TIGR00936; ahcy; 1.
CC PROSITE; PS00738; ADOHCYASE_1; 1.
CC PROSITE; PS00739; ADOHCYASE_2; 1.
CC Hydrolase; NAD; One-carbon metabolism; 3D-structure.
KW INIT MET 0 0
FT NP BIND 213 244 NAD (POTENTIAL).
FT STRAND 7 8
FT TURN 11 13
FT HELIX 14 24
FT HELIX 25 27
FT HELIX 29 38
FT TURN 39 41
FT TURN 43 46
FT STRAND 48 52
FT HELIX 57 68
FT TURN 69 70
FT STRAND 72 76
FT TURN 80 82
FT HELIX 85 93
FT TURN 94 95
FT STRAND 98 100
FT TURN 102 103
FT HELIX 106 115
FT TURN 116 117
FT STRAND 118 119
FT TURN 120 121
FT STRAND 122 123
FT STRAND 126 129
FT HELIX 133 141
FT HELIX 143 146
FT TURN 147 148
FT STRAND 151 154
FT HELIX 157 168
FT TURN 169 170
FT STRAND 176 178
FT TURN 180 181
FT HELIX 183 189
FT TURN 190 190
FT HELIX 191 206
FT TURN 211 212
FT STRAND 214 218
FT HELIX 222 233
FT TURN 234 235
FT STRAND 237 241
FT HELIX 245 252
FT TURN 253 255
FT STRAND 257 259
FT HELIX 261 264
FT TURN 265 267
FT STRAND 270 273
FT STRAND 281 281
FT HELIX 283 286
FT TURN 287 288
FT TURN 291 292
FT STRAND 293 297
FT TURN 301 303
FT STRAND 305 305
FT HELIX 307 313
FT STRAND 314 321

```

FT TURN 322 323
FT STRAND 324 329
FT TURN 330 331
FT STRAND 332 338
FT TURN 333 338
FT STRAND 334 341
FT TURN 335 341
FT STRAND 336 343
FT TURN 337 343
FT STRAND 338 348
FT TURN 339 348
FT STRAND 340 354
FT TURN 341 354
FT STRAND 342 373
FT TURN 343 373
FT STRAND 344 384
FT TURN 345 384
FT STRAND 346 397
FT TURN 347 397
FT STRAND 348 401
FT TURN 349 401
FT STRAND 350 416
FT TURN 351 416
FT STRAND 352 428
FT TURN 353 428
SQ SEQUENCE 431 AA; 47407 MW; 537154C9EAA4E380 CRC64;

Query Match
Best Local Similarity 96.2%; Score 2184; DB 1; Length 431;
Matches 415; Conservative 96.3%; Pred. No. 1.8e-156;
Mismatches 7; Indels 0; Gaps 0;

QY 2 SDKLPYKADIGLAAGWKALDIAENMPGLMRERYASAKPLKGARIAGCLHMTVETA 61
DB 1 ADKLPYKADIGLAAGWKALDIAENMPGLMRERYASAKPLKGARIAGCLHMTVETA 60

QY 62 VLTIETLVTLGAEVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLMWCIETQTLVF 121
DB 61 VLTIETLVTLGAEVQWSSCNIFSTQDHAAAAIAKAGIPVYAWKGETDEEYLMWCIETQTLVF 120

QY 122 DGPLNMLDDGGDLTLNLIHTKYPQLPGIRGISEETTTGVHNLKXMGANGILKVPAINVN 181
DB 121 DGPLNMLDDGGDLTLNLIHTKYPQLPGIRGISEETTTGVHNLKXMGANGILKVPAINVN 180

QY 182 DSVTKSKFDNLYGCRSLIDGIRKATDVMITAGVAVVAGYGVKGCAQALRGFGARV 241
DB 181 DSVTKSKFDNLYGCRSLIDGIRKATDVMITAGVAVVAGYGVKGCAQALRGFGARV 240

QY 242 TEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHPEQMKDDAIVCNIGH 301
DB 241 TEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHPEQMKDDAIVCNIGH 300

QY 302 FDVEIDVKMLNENAVEKVNIPQVDYRLKNGRRRIILAEGRVLNVLGCAMGHPSPFVMS 361
DB 301 FDVEIDVKMLNENAVEKVNIPQVDYRLKNGRRRIILAEGRVLNVLGCAMGHPSPFVMS 360

QY 362 FTNQVMAQIELWTHPDKYPVGVHFLPKLDEAVAEAHGLKLVKLTCLKTEKQAQYLGMP 421
DB 361 FTNQVMAQIELWTHPDKYPVGVHFLPKLDEAVAEAHGLKLVKLTCLKTEKQAQYLGMP 420

QY 422 DGPFPKPDHYRY 432
DB 421 NGPFPKPDHYRY 431

RESULT 4
SAH1_XENLA
ID -SAH1_XENLA STANDARD; PRT; 433 AA.
AC P51853;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosylhomocysteine 1 (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase 1) (ADOHCYASE 1).
GN AHY1 OR AHY.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95110290; PubMed=7811234;

```

```

RA Seery L.T., McCabe B.D., Schoenberg D.R., Whitehead A.S.;
RT "S-adenosyl-L-homocysteine hydrolase from Xenopus laevis --
RL Biochem. Biophys. Res. Commun. 205:1539-1546(1994).
CC 1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC 1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC 1- COFACTOR: NAD (BY SIMILARITY).
CC 1- PATHWAY: Activated methyl cycle.
CC 1- SIMILARITY: Belongs to the adenosylhomocysteine family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; L35559; AAA65963.1; -
DR PIR; JC2480; JC2480.
DR HSSP; P10760; 1D4F.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF05221; AdoHcyase_1.
DR Pfam; PF00870; AdoHcyase_NAD; 1.
DR TIGRPFAM; TIGR00936; ahcy; 1.
DR PROSITE; PS00738; ADOHCYASE_1; 1.
DR PROSITE; PS00739; ADOHCYASE_2; 1.
KW Hydrolase; NAD; One-carbon metabolism.
FT NP BIND 215 246 NAD (POTENTIAL).
SQ SEQUENCE 433 AA; 47747 MW; 28C7DB273199F2FA CRC64;

Query Match 88.8%; Score 2017.5; DB 1; Length 433;
Best Local Similarity 87.8%; Pred. No. 5.6e-144;
Matches 380; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY 1 MSDKLPYKADIGLAAGWKALDIAENMPGLMRERYASAKPLKGARIAGCLHMTVET 60
DB 1 MSDKLSYKADISLADWGRKAIEIAENEMPGMLKQRMHSESKPLKGARIAGCLHMTLOT 60

QY 61 AVLTIETLVTLGAEVQWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEEYLMWCIETQTLVF 120
DB 61 AVLTIETLVTLGAEVQWSSCNIFSTQDHAAAAIAKAGIPVYAWKGETDEEYLMWCIETQTLVF 120

QY 121 KDG-PLNMLDDGGDLTLNLIHTKYPQLPGIRGISEETTTGVHNLKXMGANGILKVPAIN 179
DB 121 KDGKPLNMLDDGGDLTLNLIHTKYPQLPGIRGISEETTTGVHNLKXMGANGILKVPAIN 180

QY 180 VNSVTYKSKFDNLYGCRSLIDGIRKATDVMITAGVAVVAGYGVKGCAQALRGFGARV 239
DB 181 VNSVTYKSKFDNLYGCRSLIDGIRKATDVMITAGVAVVAGYGVKGCAQALRGFGARV 240

QY 240 IITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHPEQMKDDAIVCNI 299
DB 241 IITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCADIVEGRHPEQMKDDAIVCNI 300

QY 300 GHFDVEIDVKMLNENAVEKVNIPQVDYRLKNGRRRIILAEGRVLNVLGCAMGHPSPFVMS 359
DB 301 GHFDVEIDVKMLNENAVEKVNIPQVDYRLKNGRRRIILAEGRVLNVLGCAMGHPSPFVMS 360

QY 360 NSFTNQVMAQIELWTHPDKYPVGVHFLPKLDEAVAEAHGLKLVKLTCLKTEKQAQYLG 419
DB 361 NSFTNQVMAQIELWTHPDKYPVGVHFLPKLDEAVAEAHGLKLVKLTCLKTEKQAQYLG 420

QY 420 SCDBGPKPDHYRY 432
DB 421 DKEGFPKPDHYRY 433

```

```
RESULT 5
SAH2_XENLA STANDARD; PRT; 433 AA.
AC O93477;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosylhomocysteine 2 (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase 2) (ADHCYASE 2).
GN AHCY2 OR SAHH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Radomski N., Plessmann U., Mohl C., Weber K., Dreyer C.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE. (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -!- COFACTOR: NAD (BY SIMILARITY).
CC -!- PATHWAY: Activated methyl cycle.
CC -!- SIMILARITY: Belongs to the adenosylhomocysteine family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ007835; CAA07706.1; -.
CC HSSP; P10760; 1D4F.
CC InterPro; IPR000043; Ado_hcyase.
CC Pfam; PF05221; AdoHcyase_1.
CC DR TIGRFAMs; TIGR00936; ahcy1.
CC DR PROSITE; PS00738; ADHCYASE_1; 1.
CC DR PROSITE; PS00739; ADHCYASE_2; 1.
CC KW Hydrolase; NAD; One-carbon metabolism.
CC FT NP BIND 215 246 NAD (POTENTIAL).
CC SEQUENCE 433 AA; 47745 MW; 3CB91D67C555B47C CRC64;

Query Match 98.7%; Score 2015.5; DB 1; Length 433;
Best Local Similarity 87.3%; Pred. No. 7.9e-144;
Matches 378; Conservative 29; Mismatches 25; Indels 1; Gaps 1;

QY 1 MSDKLPYKADTGLAAGKALDIAENPGLMWRERYASKPLKGARIAGLHMTVT 60
DB 1 MSDKLSYKADISLADWGRKAEIAENPGLMKRMYSESKPLKGARIAGLHMTLT 60
QY 61 AVLIETLVLTGAEVQWSSCNIFSTQNHAAAAAKAGIPVYAWKGETDEEYLWCIEOTLYF 120
DB 61 AVLIETLTGAIEVQWSSCNIFSTQDHAATAKTGVPVYAWKGETDEEYIWCIEGTIYF 120
QY 121 KQG-PLNMILDGGDLTNLIHTKYPOLLPGIRGISETTTGVNLYKMWANGILKVPAIN 179
DB 121 KDGKPLNMLDGGDLTNLVHTKYPOLLKIGKISETTTGVNLYKMKSSGTLQVPAIN 180
QY 180 VNDVSTKSFNLYGCRSLIDGKIKATDVIAGVAVAGYCDVCKGCAQALRGFGARV 239
DB 181 VNDVSTKSFNLYGCRSLIDGKIKATDVIAGVAVAGYCDVCKGCAQALRGFGARV 240
QY 240 IITEIDPINALQAAMEGYEVTTMDEACQEGNIFVTTCGIDILGRHFQMKDDAIVCNI 299

RESULT 6
SAHH_DROME STANDARD; PRT; 431 AA.
AC Q27580;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosylhomocysteine (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN AHCY13 OR AHCY.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97188525; PubMed=90371110;
RA Caggese C., Ragone G., Barsanti P., Moschetti R., Messina A.,
RA Massari S., Caizzi R.;
RT "The S-adenosyl-L-homocysteine hydrolase of Drosophila melanogaster:
RT identification, deduced amino acid sequence and cytological
RT localization of the structural gene.";
RL Mol. Gen. Genet. 253:492-498 (1997).
CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -!- COFACTOR: NAD (BY SIMILARITY).
CC -!- PATHWAY: Activated methyl cycle.
CC -!- SIMILARITY: Belongs to the adenosylhomocysteine family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X95636; CAA64892.1; -.
CC HSSP; P10760; 1B3R.
CC FlyBase; FBgn014455; Ahcy13.
CC InterPro; IPR000043; Ado_hcyase.
CC Pfam; PF05221; AdoHcyase; 1.
CC Pfam; PF06670; AdoHcyase_NAD; 1.
CC DR TIGRFAMs; TIGR00936; ahcy1.
CC DR PROSITE; PS00738; ADHCYASE_1; 1.
CC DR PROSITE; PS00739; ADHCYASE_2; 1.
CC KW Hydrolase; NAD; One-carbon metabolism.
CC FT NP BIND 214 245 NAD (POTENTIAL).
CC SEQUENCE 431 AA; 47252 MW; F883FAA7DF2D898A CRC64;

Query Match 79.1%; Score 1797; DB 1; Length 431;
```



```

Best Local Similarity 80.2%; Pred. No. 1.9e-127;
Matches 345; Conservative 27; Mismatches 56; Indels 2; Gaps 2;

QY 4 KLPYKADIGLAAGRKALDIAENEMPGLMRRERYASKPLKAGIAGCLHMTVETAVL 63
DB 3 KPSYKADISLAEMGKALIIAENEMPGLMACRKYGPSKPLKAGIITCCLHMTVQTAVL 62
QY 64 IETLVTLGAQVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIBQTLVFKDG 123
DB 63 IETLVTLGAQVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYMWCIQTLLVFPDG 122
QY 124 -PLNMILDDGGDLTNLIHTKYPQLPGIRGISETTTGVHNLKYMANGILKVPAINVND 182
DB 123 QPLNMILDDGGDLTNLVAEHPELLEIRGLSETTTGVHNLKYMFGRLGMPAINVND 182
QY 183 SVTKSKFDNLGCRSLIDGIRATDVMIAKGVAVVAGVGVGKCAQALRGFGARVIT 242
DB 183 SVTKSKFDNLGCRSLIDGIRATDVMIAKGVAVVAGVGVGKCAQALRGFGARVIT 242
QY 243 EIDPINALQAAMEGVEVTTMDEACOEENIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHF 302
DB 243 EVDPINALQAAMEGVEVTTMDEACOEENIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHF 302
QY 303 DVEIDVKWLNENAVEKVNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFVMSNSF 362
DB 303 DIEIDVWLNENAVEKVNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFVMSNSF 362
QY 363 TNQVNAQIELWTHPKDPYGVGHFLPKKLDEAVAEHLGKLNKVLTKLTKEQAOYLGMSCD 422
DB 363 TNQVNAQIELWTHPKDPYGVGHFLPKKLDEAVAEHLGKLNKVLTKLTKEQAOYLGMSCD 422
QY 423 GPFKPDHYRY 432
DB 422 GPFKPDHYRY 431

RESULT 7
SAHH ANOGA STANDARD; PRT; 432 AA.
ID SAHH ANOGA AC 076757;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G3,
RA Zhao Y., Lu W., Eggleston P.;
RT "The S-adenosyl-L-homocysteine hydrolase of Anopheles gambiae.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ADENOSYLOHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLOHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLOHOMOCYSTEINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: Activated methyl cycle.
CC -1- SIMILARITY: Belongs to the adenosylhomocysteinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL; AF080546; AAC29475.1; -.
DR HSSP; P10760; 1D4P.
DR InterPro; IPR000043; Ado hcyase.
DR Pfam; PF05221; AdoHcyase_1.
DR Pfam; PF00670; AdoHcyase_NAD; 1.
DR TIGRfam; TIGR00936; arcY; 1.
DR PROSITE; PS00738; ADHOCYASE 1; 1.
DR PROSITE; PS00739; ADHOCYASE 2; 1.
KW Hydrolase; NAD; One-carbon metabolism.
FT NP_BIND 214 245 NAD (POTENTIAL).
SQ SEQUENCE 432 AA; 47741 MW; B3401186D41A3C66 CRC64;

Query Match 77.8%; Score 1767.5; DB 1; Length 432;
Best Local Similarity 77.7%; Pred. No. 3.1e-125;
Matches 334; Conservative 39; Mismatches 56; Indels 1; Gaps 1;

QY 4 KLPYKADIGLAAGRKALDIAENEMPGLMRRERYASKPLKAGIAGCLHMTVETAVL 63
DB 3 KPSYKADISLAEMGKALIIAENEMPGLMACRKYGPSKPLKAGIITCCLHMTIQTAVL 62
QY 64 IETLVTLGAQVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIBQTLVFKDG 123
DB 63 IETLVTLGAQVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYMWCIQTLLVFPDG 122
QY 124 -PLNMILDDGGDLTNLIHTKYPQLPGIRGISETTTGVHNLKYMANGILKVPAINVND 182
DB 123 KPLNMILDDGGDLTNLVAEHPELLEIRGLSETTTGVHNLKYMFGRLGMPAINVND 182
QY 183 SVTKSKFDNLGCRSLIDGIRATDVMIAKGVAVVAGVGVGKCAQALRGFGARVIT 242
DB 183 SVTKSKFDNLGCRSLIDGIRATDVMIAKGVAVVAGVGVGKCAQALRGSGRVLIT 242
QY 243 EIDPINALQAAMEGVEVTTMDEACOEENIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHF 302
DB 243 EIDPINALQAAMEGVEVTTMDEACOEENIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHF 302
QY 303 DVEIDVKWLNENAVEKVNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFVMSNSF 362
DB 303 DCEINVTWLNENAVEKVNIPQVDYRLANGNHIILLAEGRVLNLCGAMGHPSPFVMSNSF 362
QY 363 TNQVNAQIELWTHPKDPYGVGHFLPKKLDEAVAEHLGKLNKVLTKLTKEQAOYLGMSCD 422
DB 363 TNQVNAQIELWTHPKDPYGVGHFLPKKLDEAVAEHLGKLNKVLTKLTKEQAOYLGMSCD 422
QY 423 GPFKPDHYRY 432
DB 423 GPFKPDHYRY 432

RESULT 8
SAHH CAEEL STANDARD; PRT; 437 AA.
ID SAHH CAEEL AC 27604;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodermata; Nematocera; Caenorhabditidae.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93209535; PubMed=8458573;
RA Prasad S.S., Starr T., Rose A.M.;
RT "Molecular characterization in the dpy-14 region identifies the
RT adenosylhomocysteine hydrolase gene in Caenorhabditis elegans.";
RL Genome 36:57-65(1993).
RN [2]
RP SEQUENCE FROM N.A.
```

RC STRAIN=Bristol N2;
RA Maggi L., Goela D.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Activated methyl cycle.
CC -1- SUBUNIT: Homotetramer.
CC -1- SIMILARITY: Belongs to the adenosylhomocysteinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M64306; AAA28062.1; -;
CC EMBL; S57284; AAB25906.1; -;
CC EMBL; AF043699; AAB97565.1; -;
CC PIR; T32918; T32918.
CC KSSP; P10760; ID4F.
CC WormPep; K02F2.2; CE17154.
CC InterPro; IPR000043; Ado_hcyase.
CC Pfam; PF05221; AdoHcyase; 1.
CC Pfam; PF00670; AdoHcyase_NAD; 1.
CC TIGRfam; TIGR00936; ahcY; 1.
CC PROSITE; PS00738; ADHCYASE_1; 1.
CC PROSITE; PS00739; ADHCYASE_2; 1.
CC Hydrolase; NAD; One-carbon metabolism.
CC NP BIND 216 247 NAD (POTENTIAL).
CC FT NAD
CC SQ SEQUENCE 437 AA; 47535 MW; 53ADAB24507CFCD6 CRC64;

Query Match 77.0%; Score 1748; DB 1; Length 437;
Best Local Similarity 76.7%; Pred. No. 9.3e-124;
Matches 332; Conservative 41; Mismatches 56; Indels 4; Gaps 2;

QY 4 KLPYKADVADGLAAGRWKALDIAENPGLMRWEERYASAKPLKAGRIAGLHMTVTAVL 63
DB 5 KPAVKADVADGLADFGKREIILAENPGLMAMSKYGPSOPLKAGRIAGLHMTIQTAVL 64
QY 64 IETLVLTGAEVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLCIEQTLVFKDG 123
DB 65 IETLTALGAEVQWSSCNIFSTQDHAAAIQAQTGVPVYAWKGETDEEYLCIEQTLVFKDG 124
QY 124 -PLNMLDGGDLTNLHVKYPLLPGIRGISETTGGVHNLYKMWANGILKVPAINVD 182
DB 125 QPLNMLDGGDLTNLVHAKYQPLAGIRGLSETTGGVHNLAQKGLKVPAINVD 184
QY 183 SVTKSFQNLGYCRESLIDGIKRATDVMTAGVAVAGYGVKGCQAALRGFGARVIT 242
DB 185 SVTKSFQNLGYRESLPGIKRATDVMTAGVAVAGYGVKGCQAALRGFGARVIT 244
QY 243 EIDPINALQAAMEGYVTMTDEACQENIFVTTTGCIDILGRHFFQMKDDAIVCNIGHF 302
DB 245 EIDPINALQAAMEGYVTMTLEEAAPKANIIVTTGCKDIVTKGFELLPNDALVNVGHF 304
QY 303 DVEIDVKWLNENAVEKNIKPQVDRLKNGRIILAEGLVNLGCAMGHPFSFVMSNSF 362
DB 305 DCEIDVKWLNATKKTIDKQVDRLKNGRIILAEGLVNLGCAMGHPFSFVMSNSF 364
QY 363 TNQVMAQIELWTH---PDKYPGVGHFLPKKLDEAAVEAHGLKLVNKLTKLTKQAOYLGM 419
DB 365 TNQVLQVLEWTKFGTPQYKGLVLPKTLDEEVAYLHLAQGLVNLKLSDEQAQYLG 424
QY 420 SCDGPFKPDHYRY 432
||:|||||

Db 425 PVAGPYKPDHYRY 437

RESULT 9
SAHH_SCHPO
ID SAHH_SCHPO STANDARD; PRT; 433 AA.
AC O13639;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN SPBC8D2.18C OR P1047.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetaceae;
OC Schizosaccharomycetales; Schizosaccharomycetes;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=20089027; PubMed=10620777;
RA Machida M., Yamazaki S., Kunihiro S., Tanaka T., Kushida N., Jinno K.,
RA Haikawa Y., Yamazaki J., Yamamoto S., Sekine M., Oguchi A., Nagai Y.,
RA Sakai M., Aoki K., Ogura K., Kudoh Y., Kikuchi H., Zhang M.Q.,
RA Yanagida M.;
RT "A 38 kb segment containing the cdc2 gene from the left arm of fission
RT yeast chromosome II: sequence analysis and characterization of the
RT genomic DNA and cDNAs encoded on the segment.";
RL Yeast 16:71-80(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Ostell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Welljens I., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Activated methyl cycle.
CC -1- SIMILARITY: Belongs to the adenosylhomocysteinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

```

CC -----
DR EMBL; AB004537; BRA21427.1; -.
DR EMBL; AL020272; CAA17833.1; -.
DR PIR; T40763; T40763.
DR HSSP; P10760; 183R.
DR GeneDB_Spombe; SPBC8D2.18c; -.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF05221; AdoHcyase; 1.
DR TIGRfams; TIGR00936; ahcy; 1.
DR PROSITE; PS00738; ADHCYASE_1; 1.
DR PROSITE; PS00739; ADHCYASE_2; 1.
KW Hydrolyase; NAD; One-carbon metabolism.
FT NP_BIND 215 246 NAD (POTENTIAL).
SQ SEQUENCE 433 AA; 47383 MW; 8B4279CC2B518E13 CRC64;

Query Match 76.3%; Score 1733; DB 1; Length 433;
Best Local Similarity 77.4%; Pred. No. 1.2e-122;
Matches 330; Conservative 40; Mismatches 56; Indels 2; Gaps 2;

QY 7 YKVADIGLAAGRKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVETAVLIET 66
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 YKVADISLAAGRKALDIAENEMPGLIANREKAKSQPLKGARIAGCLHMTIQTAVLIET 65
QY 67 LVTLAGVQWSSCNIFSTQNHAAAAIAKAGIPYAWKGETDEBYLWCIEQTL-YFKDG-P 124
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 LVALGAEVWSSCNIVSTQDHAAAAIAATGVPVPAWKGETEYBYLWCIEQQLKSPFGKP 125
QY 125 LNMILDDGGDLTLNLIHTKYPQLLPGIRGISEETTTGVNLYKMGANGILKVPAINVDSV 184
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 LNMILDDGGDLTLNLIHTKYPQLLPGIRGISEETTTGVNLYKMGANGILKVPAINVDSV 185
QY 185 TSKSFNLYGCRSLIDGIRKATDVMIAGKAVVAGVGVGKCAQALRGFGARVLIITEI 244
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 TSKSFNLYGCRSLIDGIRKATDVMIAGKAVVAGVGVGKCAQALRGFGARVLIITEI 245
QY 245 DPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIILGRHFQMKDDAIVCNIGHFDV 304
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 DPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIILGRHFQMKDDAIVCNIGHFDV 305
QY 305 EIDVKVLENAVEKVNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFVMSGFTN 364
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 EIDVAVLKNANADVNNIPQVDYRLKNGRRIILLAEGRVLNLCGAMGHPSPFVMSGFTN 365
QY 365 QVMAQIELWTHPDYKPVGVHFLPKKLEAVABAHGLKLVNKLTKLEKQAQYLGMSCDGP 424
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366 QVLAQIALWTDNTSYPLGVNMLPKLDEEVARHLGLKGLVLTLSVQSDYLGIPVDGP 425
QY 425 FKPDHYRY 432
DB :|||||
426 YKADHYRY 433

RESULT 10
SAHH_DICDI STANDARD; PRT; 430 AA.
AC P10819;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosylhomocysteine (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN SAHA.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88240426; PubMed=3288206;
RA Kasir J., Aksamit R.R., Backlund P.S. Jr., Cantoni G.L.;
RT "Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from
RT Dictyostellium discoideum as deduced from the cDNA sequence.";
RL Biochem. Biophys. Res. Commun. 153:359-364(1988).

```

```

[2]
RN RP SEQUENCE OF 258-430 FROM N.A.
RX MEDLINE=89000991; PubMed=3139100;
RA Guitton M.C., Part D., Veron M.;
RT "Cloning of a cDNA for the S-adenosyl-L-homocysteine hydrolase from
RT Dictyostellium discoideum.";
RL Biochimie 70:835-840(1988).
CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Activated methyl cycle.
CC -1- SUBUNIT: Homotetramer.
CC -1- SIMILARITY: Belongs to the adenosylhomocysteine family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19937; AAA33165.1; -.
DR EMBL; X12523; CAA31040.1; -.
DR PIR; A27655; A27655.
DR HSSP; P10760; 1D4F.
DR DictyDb; D005004; saba.
DR InterPro; IPR000043; Ado_hcyase.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF05221; AdoHcyase; 1.
DR TIGRfams; TIGR00936; ahcy; 1.
DR PROSITE; PS00738; ADHCYASE_1; 1.
DR PROSITE; PS00739; ADHCYASE_2; 1.
KW Hydrolyase; NAD; One-carbon metabolism.
FT NP_BIND 213 244 NAD (POTENTIAL).
FT CONFLICT 258 258 I -> V (IN REF. 2).
FT CONFLICT 261 265 METAA -> KSSPWKP (IN REF. 2).
FT CONFLICT 268 269 SN -> HYOT (IN REF. 2).
FT CONFLICT 312 312 A -> G (IN REF. 2).
FT CONFLICT 324 324 D -> S (IN REF. 2).
FT CONFLICT 360 364 FCNQT -> SVTK (IN REF. 2).
FT CONFLICT 383 383 L -> P (IN REF. 2).
SQ SEQUENCE 430 AA; 47083 MW; D56DF0F84DE45774 CRC64;

Query Match 74.7%; Score 1697.5; DB 1; Length 430;
Best Local Similarity 74.6%; Pred. No. 5.6e-120;
Matches 320; Conservative 38; Mismatches 70; Indels 1; Gaps 1;

QY 4 KLPYKVADIGLAAGRKALDIAENEMPGLMRMRYSASKPLKGARIAGCLHMTVETAVL 63
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 KLPYKVADISLAAGRKALDIAENEMPGLMTRKKYGPQALKGARIAGCLHMTIQTAVL 62
QY 64 IETLVTLGAEVQWSSCNIFSTQNHAAAAIAKAGIPYAWKGETDEBYLWCIEQTLYPQDG 123
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 IETLVTLGAEVQWSSCNIFSTQDHAAAAIAATGVPVPAWKGETEYBYLWCIEQTLVQDG 122
QY 124 PLNMILDDGGDLTLNLIHTKYPQLLPGIRGISEETTTGVNLYKMGANGILKVPAINVDS 193
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 PLNMILDDGGDLTLNLIHTKYPQLLPGIRGISEETTTGVNLYKMGANGILKVPAINVDS 192
QY 184 VTKSKFDNLYGCRSLIDGIRKATDVMIAGKAVVAGVGVGKCAQALRGFGARVLIITE 243
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
183 VTKSKFDNLYGCRSLIDGIRKATDVMIAGKAVVAGVGVGKCAQALRGFGARVLIITE 242
QY 244 IDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIILGRHFQMKDDAIVCNIGHFD 303
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
243 IDPINALQAAMEGYEVTMTDEACQEGNIFVTTTGCIDIILGRHFQMKDDAIVCNIGHFD 302

```

```

304 VEIDVKWLNENAVEKVNIPQVDYRLKXGRRILLIAEGRVLNLCGAMGHPSPFMSNST 363
303 CEIDVAWLANA-KDVTVPQVDYRLANGVHILLAEGRVLNLCGCTGHPSPFMSNSTC 361
364 NQMAQIELWTHDPKYPVGVHPLPKLDEABAHLGKLVNLTKEQAQYLGMSCDG 423
362 NQTLAQIALWTKTEEYPLGVHLLPKLIDEEVARLHLDLQAKLTTTEKQSEVLSVPVAG 421
424 PKPDHYRY 432
422 PYKVDHYRY 430

BEST LOCAL SIMILARITY 70.8%; Pred. No. 4e-115;
Matches 313; Conservative 52; Mismatches 61; Indels 16; Gaps 3;

QY 7 YKADIIGLAAGRKALDIAENEMPGIMRREYSASKPLKGARIACCIHMTVETAVLIET 66
DB 8 YKIADISLAAGRKETEIAEHEMPLMAIRKAYGDVQPLKGARIACCLHMTITQTAVLIET 67
QY 67 LVTLAGAEVOWSCNTFSTONHAAAIKAGIPYAWKGTDEEYLWCISOTLY-FKDG-P 124
DB 68 LVALGAEVWSSCNYSYTDHAAAIASGVVPFAWKGETEELWCISQQLFAFDKNNK 127
QY 125 LNMILDGCGDLNLNHTKYPOLLPGIRGISEETTTGVHNLKYMANGILKVPAINVDSV 184
DB 128 LNLILDGCGDLTLVHEKHPEMLEDCFGLSEETTTGVHLYRMVKGKLVPAINVDSV 187
QY 185 TKSKFDNLYGCRSLIDGIKRATDVMIAKVAVAGYGVGKGAQALRGFGARVITTEI 244
DB 188 TKSKFDNLYGCRSLVDGIKRATDVMIAKVAVAGYGVGKGAQALRGFGARVITTEI 247
QY 245 DPINALQAAMEGYEVTMTDEACQEGNI FVTTTGCIDIIILGRHFEQMKDDAIVCNIGHFDV 304
DB 248 DPINALQAAMEGYQVVTMEDASHIGQVFTTTGCRDIINGEHFINNPEDAIVCNIGHFDI 307
QY 305 EIDVKWLNENAVEKVNIPQVDYRLKXGRRILLIAEGRVLNLCGAMGHPSPFMSNSTN 364
DB 308 EIDVAWLKANAKECINIKPQVDYRLSSGRHVLILLANGRLVNLGCATGHSSPFMSNSTN 367
QY 365 QVMAQIELWTHDPK-----YPVGVHPLPKLDEABAHLGKLVNLTKE 410
DB 368 QVLAQIALFKSNDKSFREKHIEFQKTGPPEVGVHPLPKLIDEEVARLHLDLQAKL 427
QY 411 EKOQAYLGMSCDGPFKPDHYRY 432
DB 428 KVQSEYLGIPREGPFKADHYRY 449

RESULT 12
SAHH_YEAST
ID SAHH_LEIDO STANDARD; PRT; 437 AA.
AC P36889;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
OS Leishmania donovani.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5661;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92365726; PubMed=1501636;
RA Henderson D.M., Hanson S., Allen T., Wilson K., Coulter-Karis D.E.,
RA Greenberg M.L., Hersfield M.S., Ullman B.;
RT "Cloning of the gene encoding Leishmania donovani
RT S-adenosylhomocysteine hydrolase, a potential target for
RT antiparasitic chemotherapy.";
RL Mol. Biochem. Parasitol. 53:169-183 (1992).
CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: Activated methyl cycle.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the adenosylhomocysteinase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U18796; AAB64578.1; --
CC DR PIR; S50546; S50546.
CC DR HSP; P10760; I83R.
CC DR SGD; S000845; SAH1.
CC DR GO; GO:0004013; P:adenosylhomocysteinase activity; NAS.
CC DR GO; GO:0006555; P:methionine metabolism; NAS.
CC DR GO; GO:0016259; P:selenocysteine metabolism; NAS.
CC DR InterPro; IPR000043; Ado_hcyase.
CC DR Pfam; PF05221; AdoHcyase; 1.
CC DR Pfam; PF00670; AdoHcyase_NAD; 1.
CC DR TIGRFAMs; TIGR00936; ahcy; 1.
CC DR PROSITE; PS00738; ADOHCYASE_1; 1.
CC DR PROSITE; PS00739; ADOHCYASE_2; 1.
CC DR Hydrolase; NAD; One-carbon metabolism.
CC NP_BIND 217 248 NAD (POTENTIAL).
CC SQ SEQUENCE 449 AA; 49125 MW; D98D9DD329374F5B CRC64;

Query Match 71.9%; Score 1633; DB 1; Length 449;

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR HMBL; M76556; AAA29265.1; -
 DR HSSP; P10760; 1B3R.
 DR InterPro; IPR000043; Ado hcyase.
 DR Pfam; PF05221; AdoHcyase; 1.
 DR Pfam; PF06670; AdoHcyase; NAD; 1.
 DR TIGRfam; TIGR00936; ahcy; 1.
 DR PROSITE; PS00738; ADOHCYASE_1; 1.
 DR PROSITE; PS00739; ADOHCYASE_2; 1.
 KW Hydrolyase; NAD; One-carbon metabolism.
 FT NP_BIND 213 244 NAD (POTENTIAL).
 SQ SEQUENCE 437 AA; 47791 MW; D687C918DE2B0637 CRC64;

Query Match 70.8%; Score 1607.5; DB 1; Length 437;
 Best Local Similarity 70.5%; Pred. No. 3.2e-113;
 Matches 310; Conservative 46; Mismatches 73; Indels 11; Gaps 4;

QY 1 MSKLPYKADVGLAAGKRGKALDIAENEMPGLMRMRERYASAKPLKGARIAGCLHMTVET 60
 DB 1 MAD---YKVKDLSLAEWGRKALIAENEMPGLMRLREYGPSQPLKGARIAGCLHMTVQT 57
 QY 61 AVLIETLVTLGAEVOWSSCNIFSTONHAAAIAKAGIPVYANKGETDEEYLWCIEQTLL-- 118
 DB 58 AVLIETLKALGADVWSSCNIFSTONHAAAIAKAGVPVFAWKGETDEEYEWICIAQTVGK 117
 QY 119 YPKDGLNMLDDGDLTLNHTKYPQLPGIRGISEETTTGVNLYKMWANGILKVPAL 178
 DB 118 PSGDGLPNMLDDGDLTLNHTKYPQLPGIRGISEETTTGVNLYKELSGKLNPMCAI 177
 QY 179 NVNDSVTSKFPNLVGCRESLIDGIRKATDVMIAKGVAVVAGVGVGKCAQALRGFGAR 238
 DB 178 NVNDSVTSKFPNLVGCRESLIDGIRKATDVMIAKGVAVVAGVGVGKCAQALRAFGAR 237
 QY 239 VIITEIDPINALQAAMEGVEVTTMDACOEENIFVTTTCIDIIILGRHFEQMKDDAIVCN 298
 DB 238 VVTEVDPIALQAAMEGQVALVEDVMADAHIFVTTTGNDDIIITSEHPHMRDDAIVCN 297
 QY 299 IGHFDVEIDVKWLENAVEKNVKNPKQVDYRRLKNGRRRIILLAEGRVLNLCGAMGHPSPFM 358
 DB 298 IGHFDTEIQVGLWLEANEKEHVEIKQVDYRWENGRIILLAKRVLNLCGASGHPSPFM 357
 QY 359 SNSFTNQVMAQIELWTHPD--KYP-----VGVHPLPKLDEAVAHAHLGKLVNKLTKLTEK 412
 DB 358 SNSFTNQVLAQIELWTHPD--KYP-----VGVHPLPKLDEAVAHAHLGKLVNKLTKLTEK 411
 QY 413 QAQYLGMSCDGPFKPDHYRY 432
 DB 418 QAIEYINCPVNGPFKPDHYRY 437

RESULT 13
 ID SAHH_PNECA STANDARD; PRT; 440 AA.
 AC Q12663;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenosylhomocysteinease (EC 3.3.1.1) (S-adenosyl-L-homocysteine
 DE hydrolyase) (AdoHcyase) (Fragment).
 GN SAHH.
 OS Pneumocystis carinii.
 OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
 OC Pneumocystis.
 OX NCBI_TaxID=4754;
 RP [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Rattus;
 RX MEDLINE=96420115; PubMed=8822814;

RA Lasbury M.E., Brady S., McLaughlin G., Bartlett M.S., Smith J.W.,
 RA Lee C.H.;
 RT "Cloning of the S-adenosylhomocysteine hydrolase gene of *Pneumocystis*
 RT *carinii*";
 RL J. Eukaryot. Microbiol. 43:68-6S(1996).
 CC -!- FUNCTION: ADENOSYLOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
 CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
 CC THEREFORE ADENOSYLOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
 CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
 CC CONCENTRATION OF ADENOSYLOMOCYSTEINE.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
 CC + L-homocysteine.
 CC -!- COFACTOR: NAD (BY SIMILARITY).
 CC -!- PATHWAY: Activated methyl cycle.
 CC -!- SIMILARITY: Belongs to the adenosylhomocysteinease family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; U57795; AAC98514.1; -
 DR HSSP; P10760; 1B3R.
 DR InterPro; IPR000043; Ado hcyase.
 DR Pfam; PF05221; AdoHcyase; 1.
 DR Pfam; PF06670; AdoHcyase; NAD; 1.
 DR TIGRfam; TIGR00936; ahcy; 1.
 DR PROSITE; PS00738; ADOHCYASE_1; 1.
 DR PROSITE; PS00739; ADOHCYASE_2; 1.
 KW Hydrolyase; NAD; One-carbon metabolism.
 FT NON_TER 1 1
 FT NP_BIND 222 253 NAD (POTENTIAL).
 SQ SEQUENCE 440 AA; 48766 MW; BEF5360C8D9F514D CRC64;

Query Match 67.6%; Score 1536; DB 1; Length 440;
 Best Local Similarity 69.9%; Pred. No. 7.4e-108;
 Matches 306; Conservative 38; Mismatches 72; Indels 22; Gaps 7;

QY 14 LAAGRKALDIAENEMPGLMRMRERYASAKPLKGARIAGCLHMTVETAVLIETLVTLGAE 73
 DB 6 LAEWGRKALIAENEMPGLSICR-RNMRFPKPLR-VQIAGCLHMTIQTAVLIETLVTLGAE 63
 QY 74 VQWSSCNIFSTONHAAAIAKAGIPVYANKGETDEEYLWCIE-QTLYPKDGP-LNMILDD 131
 DB 64 VQWSSCNIFSTODHAAAIAAAGISVPFAWKGETDEEYLWCIEQTLYTSKGRHLMILDD 123
 QY 132 GGDNLNHTKYPQLPGIRGISEETTTGVNLYKMWANGILKVPAINVDSVTSKFPD 191
 DB 124 GGDVTSLVNHTKYPQLPGIRGISEETTTGVNLYKMWANGILKVPAINVDSVTSKFPD 183
 QY 192 LYG-----CRESLIDGIRKATDVMIAKGVAVVAGVGVGKCAQALRGF 235
 DB 184 LYGVSLVLIYKIISNLIKLRCSLIDGIRKATDVMIAKVAIVAGVGVGKCAQALRGF 243
 QY 236 GARVITTEIDPINALQAAMEGVEVTTMDACOEENIFVTTTCIDIIILGRHFEQMKDDAI 295
 DB 244 GARVITTEIDPINALQAAMEGVEVTTMDACOEENIFVTTTCIDIIILGRHFEQMKDDAI 303
 QY 296 VCNIGHFDVEIDVKWLENAVEKNVKNPKQVDYRRLKNGRRRIILLAEGRVLNLCGAMGHP 355
 DB 304 ICNIGHFDVEIDVAVLWILKKCSISINIKQVDYRRLKNGRRRIILLAEGRVLNLCGAMGHP 363
 QY 356 FVMSNFTNQVMAQIELWTHPD-KYPVGVHPLPKLDEAVAHAHLGKLVNKLTKLTEKQA 414
 DB 364 FVMSNFTNQVLAQIELWTHPD-KYPVGVHPLPKLDEAVAHAHLGKLVNKLTKLTEKQA 413
 QY 415 QYLGMSCDGPFKPDHYRY 432
 DB 423 AYLDIPIDGPKVSEHYRY 440

```

RESULT 14
SAHH CHLTE
ID SAHH CHLTE STANDARD; PRT; 471 AA.
AC Q8KEG8;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenosylhomocysteinease (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN AHCY OR SAHH OR CT0721.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Niernan W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathavan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC -|- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -|- COFACTOR: NAD (By similarity).
CC -|- PATHWAY: Activated methyl cycle.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to the adenosylhomocysteinease family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB012843; AAM71958.1; -
CC TIGR; CT0721; -.
CC HAMAP; MF_00563; -.
CC DR InterPro; IPR000043; Ado_hcyase.
CC DR InterPro; IPR000205; NAD_binding.
CC DR Pfam; PF05221; AdoHcyase_1.
CC DR Pfam; PF00670; AdoHcyase_NAD; 1.
CC DR TIGRFAMs; TIGR00936; anc7; 1.
CC DR PROSITE; PS00738; ADHCYASE_1; 1.
CC DR PROSITE; PS00739; ADHCYASE_2; 1.
CC KW Hydrolase; NAD; One-carbon metabolism; Complete proteome.
CC FT NP BIND 254 285 NAD (POTENTIAL)
CC SEQUENCE 471 AA; 51949 MW; 31B082405D241047 CRC64;
CC
Query Match 66.8%; Score 1516; DB 1; Length 471;
Best Local Similarity 64.2%; Pred. No. 2.6e-106;
Matches 300; Conservative 49; Mismatches 76; Indels 42; Gaps 6;
QY 5 LPYKVADIGLAWGRKALDAENEMGLMWRERYASRPLKCARAGGLHMTVETAVLI 64
DB 8 LDYKVADISLAEWGRKEIEIAEKEMGLMATRKXKYEKPLAGARTAGSLHMTIQAVLI 67
QY 65 ETILVTLAGEVQSSCNIFSTQNHAAIAAGKGPVVAWKGTEDEEYLWCIEOTLYPKD-- 122
DB 68 ETILVELGADVWRASCNIFSTQDHAAIAAAGVVPFAWKGTEDEEYWCIEOTLYPKD-- 127
QY 123 GPLNMTLDGGDITNLNHTKY-----PQLL----- 147
DB 128 GPLNMTLDGGDITNLNHTKY-----PQLL----- 146

```

```

QY 148 -PGIRGISSETTTGVHNLKYMANGILKVPAINVNDVTKSKFDNLYGCRSLDGIKRA 206
DB 187 AAGMGVSETTTGVHRLYQMKEGELLFPAINVNDVTKSKFDNLYGCRSLADGIKRA 246
QY 207 TDVMIAGKVAVVAGYGDVGKCAQALRGFGARVVIITEIDPINALQAAAMEGYEYTTMDEAC 266
DB 247 TDVMIAGKVAVVAGYGDVGKCAHSMRSYGARVIVTEIDPICALQAAAMEGFVTTMEEAV 306
QY 267 QEGNIFVTGTCIDITLGRHFDPMKDDAIVCNIGHFDVDEIDVKWLANE-NAVEKVNKPKQV 325
DB 307 KEGNIFVTATGNKDVITLDRHFDPMKDDAIVCNIGHFDVDEIDVKWLANE-NAVEKVNKPKQV 366
QY 326 DRYRLKNGRRITLLAEGRLVNLGCGHPSFVMSNSFTNQVMAQIELWTHPDKYPGVGVHF 385
DB 367 DRYVFENGNCIYLLAEGRLVNLGCGHPSFVMSNSFTNQVMAQIELWTHPDKYPGVGVHF 424
QY 386 LPRKLDDEAVAEAHGLKLVNKLTKLTKAQYQLGMSCDGPKDPDHYRY 432
DB 425 LPRKLDDEAVAEAHGLKLVNKLTKLTKAQYQLGMSCDGPKDPDHYRY 471
RESULT 15
SAHH XYLFA
ID SAHH XYLFA STANDARD; PRT; 480 AA.
AC Q9PEJ1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosylhomocysteinease (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN AHCY OR XF1037.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
SEQUENCE FROM N.A.
RX STRAIN=9a5C; PubMed=10910347;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Leite L.C.C.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.V., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.F., Terenzi M.F., Truffi D., Tsai S.M., Tshuko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -|- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -|- COFACTOR: NAD (By similarity).
CC -|- PATHWAY: Activated methyl cycle.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

```


Search completed: December 30, 2003, 10:41:05
Job time : 18 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2003, 10:39:03 ; Search time 21 Seconds
(without alignments)
1978.327 Million cell updates/sec

Title: US-10-043-787-1

Perfect score: 2271

Sequence: 1 MSKDLFYKVADIGLAAMGRK.....QAQLGMSCDGPFKPDHYR 432

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2266	99.8	432	1 A43629	adenosylhomocystei
2	2189	96.4	432	1 A26583	adenosylhomocystei
3	2017.5	88.8	433	1 JC2480	adenosylhomocystei
4	1748	77.0	437	2 T32918	adenosylhomocystei
5	1733	76.3	433	2 T40763	adenosylhomocystei
6	1697.5	74.7	430	1 A27655	adenosylhomocystei
7	1633	71.9	449	2 S50546	adenosylhomocystei
8	1599.5	70.4	437	1 A45569	adenosylhomocystei
9	1342	59.1	446	2 D82730	adenosylhomocystei
10	1323	58.3	485	2 S38379	adenosylhomocystei
11	1315.5	57.9	495	2 B70593	adenosylhomocystei
12	1312	57.8	462	1 A46035	adenosylhomocystei
13	1302.5	57.4	492	2 D87005	probable S-adenosyl
14	1302	57.3	481	2 AG3505	adenosylhomocystei
15	1299	57.2	466	2 AF2580	S-adenosylhomocyst
16	1299	57.2	466	2 D97362	adenosylhomocystei
17	1292	56.9	485	2 C71400	adenosylhomocystei
18	1279	56.3	485	2 S71621	adenosylhomocystei
19	1277	56.2	485	2 T06764	adenosylhomocystei
20	1256.5	55.3	483	2 H87280	adenosylhomocystei
21	1249.5	55.0	449	2 H83591	adenosylhomocystei
22	1215	53.5	479	1 A50400	S-adenosyl-L-homoc
23	1151	50.7	537	2 T08681	adenosylhomocystei
24	1121.5	49.4	597	1 S01302	adenosylhomocystei
25	965.5	42.5	442	2 C90224	S-adenosyl-L-homoc
26	961.5	42.3	417	2 S58193	adenosylhomocystei
27	943	41.5	418	2 E70401	probable adenosylh
28	927.5	40.8	399	2 B72649	adenosylhomocystei
29	878.5	38.7	422	2 B75064	adenosylhomocystei

ALIGNMENTS

RESULT 1

A43629 adenosylhomocysteinase (EC 3.3.1.1) - human
N:Alternate names: S-adenosyl-L-homocysteine hydrolase
C:Species: Homo sapiens (man)
C:Date: 11-Feb-1993 #sequence revision 04-Oct-1996 #text_change 07-May-1999
C:Accession: A43629, A60446; A57486; S65894
R:Coulter-Karis, D.E.; Hershfield, M.S.
Ann. Hum. Genet. 53, 169-175, 1989
A:Title: Sequence of full length cDNA for human S-adenosylhomocysteine hydrolase.
A:Reference number: A43629; MUID:90087640; PMID:2596825
A:Accession: A43629
A:Molecule type: mRNA
A:Residues: 1-432 <COU>
R:Cross-references: GB:M61831
R:Arredondo-Vega, F.X.; Charlton, J.A.; Edwards, Y.H.; Hopkinson, D.A.; Whitehouse, D.B.
Ann. Hum. Genet. 53, 157-167, 1989
A:Title: Isozyme and DNA analysis of human S-adenosyl-L-homocysteine hydrolase (AHCY).
A:Reference number: A60446; MUID:90087639; PMID:2574561
A:Accession: A60446
A:Molecule type: mRNA
A:Residues: 122-432 <ARR>
A:Note: the authors translated the codon CRT for residue 162 as Glu, TAC for residue 295
R:Yuan, C.S.; Borchardt, R.T.
J. Biol. Chem. 270, 16140-16146, 1995
A:Title: Photocrosslinking of human placental S-adenosylhomocysteine hydrolase with
A:Reference number: A57486; MUID:95332317; PMID:7608178
A:Accession: A57486
A>Status: preliminary
A:Molecule type: protein
A:Residues: 175-178, 'X', 180, 'X', 182-184, 'X', 186; 319-326, 'X' <YUA>
R:Gupta, R.A.; Yuan, C.S.; Ault-Riche, D.B.; Borchardt, R.T.
Arch. Biochem. Biophys. 319, 365-371, 1995
A:Title: Limited proteolysis of S-adenosylhomocysteine hydrolase: implications for the
A:Reference number: S65894; MUID:95305573; PMID:7786017
A:Accession: S65894
A>Status: preliminary
A:Molecule type: protein
A:Residues: 2-6; 104-108; 198-202 <GUP>
C:Genetics:
A:Gene: GDB:AHCY
A:Cross-references: GDB:118983; OMIM:180960
A:Map position: 20cen-20q13.1
C:Function:
A:Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to adeno
C:Superfamily: adenosylhomocysteinase
C:Keywords: homotetramer; NAD; thioether hydrolase
F:214-243/Region: beta-alpha-beta NAD nucleotide-binding fold
F:79,113/Active site: Cys #status predicted

Query Match 99.8%; Score 2266; DB 1; Length 432;
Best Local Similarity 99.8%; Pred. No. 2.5e-159;

Matches 431; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSDKLPYKADIGLAAGWKALDIAENEMPGLMRMRERYASAKPLKGARIACGLHMTVET 60
DB 1 MSDKLPYKADIGLAAGWKALDIAENEMPGLMRMRERYASAKPLKGARIACGLHMTVET 60
QY 61 AVLIETLVLTGAEVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIEQTLYF 120
DB 61 AVLIETLVLTGAEVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIEQTLYF 120
QY 121 KDGPLNMLDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVNHLXQMMANGILKVPAINV 180
DB 121 KDGPLNMLDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVNHLXQMMANGILKVPAINV 180
QY 181 NDSVTKSFDNLYGCRESLIDGIKRATDVMIAGKVAVAGYGVGKGAQALRGFGARVI 240
DB 181 NDSVTKSFDNLYGCRESLIDGIKRATDVMIAGKVAVAGYGVGKGAQALRGFGARVI 240
QY 241 ITEIDPINALQAAMEGYEVTTMDEACQEGNI FVTTTGCIDIILGRHFEQMKDDAIVCNIG 300
DB 241 ITEIDPINALQAAMEGYEVTTMDEACQEGNI FVTTTGCIDIILGRHFEQMKDDAIVCNIG 300
QY 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGRNLVNLGCAMGHPSPVMSN 360
DB 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGRNLVNLGCAMGHPSPVMSN 360
QY 361 SFTNQVMAQIELWTHDPKYPGVGHFLPKKLDEAVAEHLGKLNVLTKLTEKQAQYLGMS 420
DB 361 SFTNQVMAQIELWTHDPKYPGVGHFLPKKLDEAVAEHLGKLNVLTKLTEKQAQYLGMS 420
QY 421 CDGPFKPDHYRY 432
DB 421 CDGPFKPDHYRY 432

RESULT 2
A26583
adenosylhomocysteinase (EC 3.3.1.1) - rat
N:Alternate names: S-adenosyl-L-homocysteine hydrolase
C:Species: Rattus norvegicus (Norway rat)
C>Date: 19-Nov-1988 #sequence revision 04-Oct-1996 #text_change 18-Jun-1999
C:Accession: A26583; B26191; S69333
R:Ogawa, H.; Gomi, T.; Mueckler, M.M.; Fujioaka, M.; Backlund Jr., P.S.; Aksamit, R.R.; U
Proc. Natl. Acad. Sci. U.S.A. 84, 719-723, 1987
A:Title: Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from rat liver as de
A:Reference number: A26583; MUID:87118240; PMID:3027598
A:Accession: A26583
A:Molecule type: mRNA
A:Residues: 1-432 <OGA>
A:Cross-references: GB:M15185; NID:G202803; PIDN:AAA40705.1; PID:G202804
R:Gomi, T.; Ogawa, H.; Fujioaka, M.
J. Biol. Chem. 261, 13422-13425, 1986
A:Title: S-Adenosylhomocysteinase from rat liver. Amino acid sequences of the peptides c
A:Reference number: A92560; MUID:87008564; PMID:3759971
A:Accession: A26191
A:Molecule type: protein
A:Residues: 76-94 <GOM>
A:Accession: B26191
A:Molecule type: protein
A:Residues: 104-121 <GO2>
A>Note: in a manner saturably inhibited by adenosine (a product) or adenine (a competit
-79 and Cys-113, suggesting a role for these residues at the active site; however, neith
R:Merta, A.; Aksamit, R.R.; Kasit, J.; Cantoni, G.L.
Eur. J. Biochem. 229, 575-582, 1995
A:Title: The gene and pseudogenes of rat S-adenosyl-L-homocysteine hydrolase.
A:Reference number: S69333; MUID:95262723; PMID:7744082
A:Accession: S69333
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-432 <MER>
A:Cross-references: EMBL:U14937; NID:g1223843; PIDN:AAA32043.1; PID:g1185363
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1994
C:Genetics:
```

A:introns: 10/1; 73/3; 99/1; 149/1; 186/3; 256/1; 285/2; 324/3; 389/3

A:Function: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to adenosylhomocysteinase
C:Superfamily: adenosylhomocysteinase
C:Keywords: homotetramer; NAD; thioether hydrolase
F:214-243/Region: beta-alpha-beta NAD nucleotide-binding fold
F:79,113/Active site: Cys #status predicted

Query Match 96.4%; Score 2189; DB 1; Length 432;
Best Local Similarity 96.3%; Pred. No. 1.2e-153;
Matches 416; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

```
QY 1 MSDKLPYKADIGLAAGWKALDIAENEMPGLMRMRERYASAKPLKGARIACGLHMTVET 60
DB 1 MADKLPYKADIGLAAGWKALDIAENEMPGLMRMRERYASAKPLKGARIACGLHMTVET 60
QY 61 AVLIETLVLTGAEVQWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIEQTLYF 120
DB 61 AVLIETLVLTGAEVQWSSCNIFSTQNHAAAIKAGIPVFAWKGETDEEYLWCIEQTLYF 120
QY 121 KDGPLNMLDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVNHLXQMMANGILKVPAINV 180
DB 121 KDGPLNMLDDGGDLTNLIHTKYPQLLPGIRGISEETTTGVNHLXQMMANGILKVPAINV 180
QY 181 NDSVTKSFDNLYGCRESLIDGIKRATDVMIAGKVAVAGYGVGKGAQALRGFGARVI 240
DB 181 NDSVTKSFDNLYGCRESLIDGIKRATDVMIAGKVAVAGYGVGKGAQALRGFGARVI 240
QY 241 ITEIDPINALQAAMEGYEVTTMDEACQEGNI FVTTTGCIDIILGRHFEQMKDDAIVCNIG 300
DB 241 ITEIDPINALQAAMEGYEVTTMDEACQEGNI FVTTTGCVDIILGRHFEQMKDDAIVCNIG 300
QY 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGRNLVNLGCAMGHPSPVMSN 360
DB 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGRNLVNLGCAMGHPSPVMSN 360
QY 361 SFTNQVMAQIELWTHDPKYPGVGHFLPKKLDEAVAEHLGKLNVLTKLTEKQAQYLGMS 420
DB 361 SFTNQVMAQIELWTHDPKYPGVGHFLPKKLDEAVAEHLGKLNVLTKLTEKQAQYLGMP 420
QY 421 CDGPFKPDHYRY 432
DB 421 INGPFPDHYRY 432
```

RESULT 3

```
JC2480
adenosylhomocysteinase (EC 3.3.1.1) - African clawed frog
N:Alternate names: S-adenosyl-L-homocysteine hydrolase
C:Species: Xenopus laevis (African clawed frog)
C>Date: 29-Mar-1995 #sequence_revision 04-Oct-1996 #text_change 28-May-1999
C:Accession: JC2480
R:Seery, L.T.; McCabe, B.D.; Schoenberg, D.R.; Whitehead, A.S.
Biochem. Biophys. Res. Commun. 205, 1539-1546, 1994
A:Title: S-Adenosyl-L-homocysteine hydrolase from Xenopus laevis: identification, developm
A:Reference number: JC2480; MUID:95110290; PMID:7811234
A:Accession: JC2480
A:Molecule type: mRNA
A:Residues: 1-433 <SEE>
A:Cross-references: GB:L35559; NID:g558507; PIDN:AAA65963.1; PID:g558508
A:Experimental source: liver
C:Function:
adenosylhomocysteinase
C:Superfamily: adenosylhomocysteinase
C:Keywords: NAD; thioether hydrolase
F:215-244/Region: beta-alpha-beta NAD nucleotide-binding fold
F:79,113/Active site: Cys #status predicted
```

Query Match 88.8%; Score 2017.5; DB 1; Length 433;
Best Local Similarity 87.8%; Pred. No. 5.1e-141;
Matches 380; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY 1 MSDKLPYKADIGLAAGWKALDIAENEMPGLMRMRERYASAKPLKGARIACGLHMTVET 60

Db 1 MSKLSYKVADISLADGKRAIEAENEMPLMKRMHSESKPLKGARIAGCLHMTLQT 60
 QY 61 AVLIETLVLTGAEVQWSSCNIFSTONHAAAIKAGIPVYAWKGETDEEYLWCIEQTLF 120
 Db 61 AVLIETLVLTGAEVQWSSCNIFSTONHAAAIKAGIPVYAWKGETDEEYLWCIEQTLF 120
 QY 121 KDG-PLNMILDDGGDLTNLHHTKYPQLLPGIRGISEETTTGVHNLKMMANGILKVPAIN 179
 Db 121 KDGKPLNMLDDGGDLTNLVHAKYPOYLGRISLSEETTTGVHNLKMMANGILKVPAIN 180
 QY 180 VNSVTKSPDNLYGCRSLIDGIKRATDVMIAKGVAVVAGYGVGKCAQALRGFGARV 239
 Db 181 VNSVTKSPDNLYGCRSLIDGIKRATDVMIAKGVAVVAGYGVGKCAQALRGFGARV 240
 QY 240 IITEIDPINALQAMGEYVTTMDEACOBEGNIFVTTTGCIDIIILGRHPEOMKDDAIVCN 299
 Db 241 IITEIDPINALQAMGEYVTTMDEACOBEGNIFVTTTGCIDIIILGRHPEOMKDDAIVCN 300
 QY 300 GHFDVBDVKMLNENAVEKNIKPQVDYRLKNGRRIILAEGRVLNLCAMGHPSPVMS 359
 Db 301 GHFDIELDKMLNENAVEKNIKPQVDYRLKNGRRIILAEGRVLNLCAMGHPSPVMS 360
 QY 360 NSFTNQVMAQIELWTHDPKYPGVGHFLPKKLDDEAVAEHLGKLNKVLTKLTEKQAOYLGM 419
 Db 361 NSFTNQVMAQIELWTHDPKYPGVGHFLPKKLDDEAVAEHLGKLNKVLTKLTEKQAOYLGM 420
 QY 420 SCGGPKPDHRY 432
 Db 421 DKEGPKPDHRY 433

RESULT 4
 T32918
 adenosylhomocysteine (EC 3.3.1.1) - Caenorhabditis elegans
 N;Alternate names: S-adenosyl-L-homocysteine hydrolase
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
 C;Accession: T32918; T41749
 R;Maggi, L.; Goela, D.
 submitted to the EMBL Data Library, January 1998
 A;Description: The sequence of C. elegans cosmid K02F2.
 A;Reference number: Z21246
 A;Accession: T32918
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-437 <MAG>
 A;Cross-references: EMBL:AF043699; PIDN:AAB97565.1; GSPDB:GN00019; CESP:K02F2.2
 A;Experimental source: strain Bristol N2; clone K02F2
 R;Prasad, S.S.; Starr, T.V.; Rose, A.M.
 Genome 36, 57-65, 1993
 A;Title: Molecular characterization in the dpy-14 region identifies the adenosylhomocysteine to adenosyl-L-homocysteine hydrolase
 A;Reference number: Z22017; MUID:93209535; PMID:8458573
 A;Accession: T41749
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-437 <PRA>
 A;Cross-references: EMBL:S57284; PIDN:AAB25906.1
 C;Genetics:
 A;Gene: K02F2.2
 A;Map position: 1
 A;Introns: 74/3, 326/3
 C;Function:
 A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to adenosyl-L-homocysteine and methionine
 C;Superfamily: adenosylhomocysteinease
 C;Keywords: NAD; thioether hydrolase
 Query Match 77.0%; Score 1748; DB 2; Length 437;
 Best Local Similarity 76.7%; Pred. No. 3.8e-121;
 Matches 332; Conservative 41; Mismatches 56; Indels 4; Gaps 2;
 QY 4 KLPYKVADIGLAAGRWKALDIAENEMPLMKRMHSESKPLKGARIAGCLHMTVETAVL 63

Db 5 KPAYKVADIKLADPGRKEIILAENEMPLMKRMHSESKPLKGARIAGCLHMTVETAVL 64
 QY 64 IETLVLTGAEVQWSSCNIFSTONHAAAIKAGIPVYAWKGETDEEYLWCIEQTLF 123
 Db 65 IETLVLTGAEVQWSSCNIFSTONHAAAIKAGIPVYAWKGETDEEYLWCIEQTLF 124
 QY 124 -PLNMILDDGGDLTNLHHTKYPQLLPGIRGISEETTTGVHNLKMMANGILKVPAIN 182
 Db 125 QPLNMLDDGGDLTNLVHAKYPOYLGRISLSEETTTGVHNLKMMANGILKVPAIN 184
 QY 183 SVTKSPDNLYGCRSLIDGIKRATDVMIAKGVAVVAGYGVGKCAQALRGFGARV 242
 Db 185 SVTKSPDNLYGCRSLIDGIKRATDVMIAKGVAVVAGYGVGKCAQALRGFGARV 244
 QY 243 EIDPINALQAMGEYVTTMDEACOBEGNIFVTTTGCIDIIILGRHPEOMKDDAIVCN 302
 Db 245 EIDPINALQAMGEYVTTMDEACOBEGNIFVTTTGCIDIIILGRHPEOMKDDAIVCN 304
 QY 303 DYEIDVKMLNENAVEKNIKPQVDYRLKNGRRIILAEGRVLNLCAMGHPSPVMS 362
 Db 305 DCEIDVKMLNENAVEKNIKPQVDYRLKNGRRIILAEGRVLNLCAMGHPSPVMS 364
 QY 363 TNQVMAQIELWTHDPKYPGVGHFLPKKLDDEAVAEHLGKLNKVLTKLTEKQAOYLGM 419
 Db 365 TNQVMAQIELWTHDPKYPGVGHFLPKKLDDEAVAEHLGKLNKVLTKLTEKQAOYLGM 424
 QY 420 SCGGPKPDHRY 432
 Db 425 PVAGPKPDHRY 437

RESULT 5
 T40763
 adenosylhomocysteine - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C;Accession: T40763
 R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
 submitted to the EMBL Data Library, March 1998
 A;Reference number: Z21948
 A;Accession: T40763
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-433 <LYN>
 A;Cross-references: EMBL:AL022072; PIDN:CAA17833.1; GSPDB:GN00067; SPDB:SPBC8D2.18c
 A;Experimental source: strain 972h-; cosmid c8D2
 C;Genetics:
 A;Gene: SPDB:SPBC8D2.18c
 A;Map position: 2
 C;Superfamily: adenosylhomocysteinease
 Query Match 76.3%; Score 1733; DB 2; Length 433;
 Best Local Similarity 77.1%; Pred. No. 4.8e-120;
 Matches 330; Conservative 40; Mismatches 56; Indels 2; Gaps 2;
 QY 7 YKVADIGLAAGRWKALDIAENEMPLMKRMHSESKPLKGARIAGCLHMTVETAVL 66
 Db 6 YKVADISLAAPGRKEIILAENEMPLMKRMHSESKPLKGARIAGCLHMTVETAVL 65
 QY 67 LVTGAEVQWSSCNIFSTONHAAAIKAGIPVYAWKGETDEEYLWCIEQTLF 124
 Db 66 LVALGAEVQWSSCNIFSTONHAAAIKAGIPVYAWKGETDEEYLWCIEQTLF 125
 QY 125 LNMILDDGGDLTNLHHTKYPQLLPGIRGISEETTTGVHNLKMMANGILKVPAIN 184
 Db 126 LNMILDDGGDLTNLVHAKYPOYLGRISLSEETTTGVHNLKMMANGILKVPAIN 185
 QY 185 TSXKPDNLGCRSLIDGIKRATDVMIAKGVAVVAGYGVGKCAQALRGFGARV 244
 Db 186 TSXKPDNLGCRSLIDGIKRATDVMIAKGVAVVAGYGVGKCAQALRGFGARV 245
 QY 245 DPINALQAMGEYVTTMDEACOBEGNIFVTTTGCIDIIILGRHPEOMKDDAIVCN 304

Db 246 DPINALQAAMGDFEVTTMBEAVKEGQIFVTTTTCGRDIIRGEHFNEMKEDSIVCNIGHFDV 305
 QY 305 EIDVKWLNENAVEKNIKPQVDYRLKNGRRITLLAEGRLVNLGCGAMGHPSPVMSNSFTN 364
 Db 306 EIDVAVLKANAKADVNNIKPQVDYRLKNGRRITLLAEGRLVNLGCGATGHPSPVMSNSFTN 365
 QY 365 QVMAQIELWTHPDYKVPVGVHFLPKKLDDEAVAEHLGKLNVLTKLTKTEKQAOYLGMSCDGP 424
 Db 366 QVLAQIALMTDNTSYPLGVHMLPKKLDDEAVARHLGKLVKLTLLTSVQSDYLGIPVDGP 425
 QY 425 FKPDHYRY 432
 Db 426 YKADHYRY 433

RESULT 6

adenosylhomocysteinase (EC 3.3.1.1) - slime mold (Dictyostelium discoideum)
 N:Alternate names: S-adenosyl-L-homocysteinase hydrolase
 C:Species: Dictyostelium discoideum
 C:Date: 31-Mar-1989 #sequence_revision 04-Oct-1996 #text_change 18-Jun-1999
 C:Accession: S06394
 R:Kasir, J.; Akemait, R.R.; Backlund Jr., P.S.; Cantoni, G.L.
 Biochem. Biophys. Res. Commun. 153, 359-364, 1988
 A:Title: Amino acid sequence of S-adenosyl-L-homocysteinase hydrolase from Dictyostelium d
 A:Reference number: A27655; MUID:88240426; PMID:3288206
 A:Accession: A27655
 A:Molecule type: mRNA
 A:Residues: 1-430 <XAS>
 A:Cross-references: GB:M1937; NID:g167663; PIDN:AAA33165.1; PID:g167664
 R:Guillon, M.C.; Part, D.; Veron, M.
 Biochimie 70, 835-840, 1988
 A:Title: Cloning of a cDNA for the S-adenosyl-L-homocysteinase hydrolase from Dictyostelium
 A:Reference number: S06394; MUID:89000991; PMID:3139100
 A:Accession: S06394
 A:Molecule type: mRNA
 A:Residues: 'VWTKSPWKPPLHYQT', 270-311, 'G', 313-323, 'S', 325-359, 'SVTK', 365-382, 'F', 384-43
 A:Cross-references: EMBL:X12523; NID:g7204; PIDN:CAA31040.1; PID:g7205
 A:Note: this sequence differs at the amino end by an apparent frame shift; comparison to
 C:Function:
 A:Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteinase to adenos
 C:Superfamily: adenosylhomocysteinase
 C:Keywords: NAD; thioether hydrolase
 F:213-242/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:78,112/Active site: Cys #status predicted

Query Match 74.7%; Score 1697.5; DB 1; Length 430;
 Best Local Similarity 74.6%; Pred. No. 2e-117;
 Matches 320; Conservative 38; Mismatches 70; Indels 1; Gaps 1;

QY 4 KLPYKVDIGLAAGKGRKALDIAENEMPGLMRERYASAKPLKGARIAGCLHMTVETAVL 63
 Db 3 KLUHYKVYDLSLAAGRKETIAENEMPGMLTKRKYGPAQILKGARIAGCLHMTTQTAVL 62
 QY 64 IETLVTLGAQVWSSCNIFSTQDQAAAAAIAAGIPVYAWKGTDEEYLWCIEQTLVFKDG 123
 Db 63 IETLVTLGAQVWSSCNIFSTQDQAAAAAIAAGIPVYAWKGTDEEYLWCIEQTLVFKDG 122
 QY 124 PLNMILDDGGDLTNLHITKYPQLPGIRGISETTTGVHNLKYMANGILKVPAINVNDV 183
 Db 123 PLNMILDDGGDLTNLHITKYPQLPGIRGISETTTGVHNLKYMANGILKVPAINVNDV 182
 QY 184 VTKSKFDNLGCRSLIDGIKRATDVMIAKGVAVAGYGDVGKCAQALRGFGARVITE 243
 Db 183 VTKSKFDNLGCRSLIDGIKRATDVMIAKGVAVAGYGDVGKCAQALRGFGARVITE 242
 QY 244 IDPINALQAMGEGYVTTWDEACQEGNIFVTTTGCIDIIILGRHFEOMKDDAIVCNIGHFD 303
 Db 243 IDPINALQAMGEGYVTTWDEACQEGNIFVTTTGCIDIIILGRHFEOMKDDAIVCNIGHFD 302
 QY 304 VIDVKWLNENAVEKNIKPQVDYRLKNGRRITLLAEGRLVNLGCGAMGHPSPVMSNSFT 363
 Db 303 CBIDVAVLNLANA-KKQTVKPVQVDYRLKNGRRITLLAEGRLVNLGCGTGHPSFVMSNSFC 361

QY 364 NOVMAQIELWTHPDYKVPVGVHFLPKKLDDEAVAEHLGKLNVLTKLTKTEKQAOYLGMSCDG 423
 Db 362 NTLAQIALMTWTESEYPLGVHLLPKKLDDEAVARHLGKLVKLTLLTKTEKQSEYLSVPVAG 421
 QY 424 FKPDHYRY 432
 Db 422 PYKVDHYRY 430

RESULT 7

S05046
 adenosylhomocysteinase (EC 3.3.1.1) - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YER043c; S-adenosyl-L-homocysteinase hydrolase
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 23-Mar-2001
 C:Accession: S05046
 R:Dieterich, F.S.
 submitted to the EMBL Data Library, December 1994
 A:Description: The sequence of S. cerevisiae cosmid 9379, 9581, and lambda clone 4678.
 A:Reference number: S050432
 A:Accession: S05046
 A:Molecule type: DNA
 A:Residues: 1-449 <DIE>
 A:Cross-references: EMBL:U18796; NID:g603265; PIDN:AAB64578.1; PID:g603276; GSPDB:GN00000
 C:Genetics:
 A:Gene: SGD:SAH1; MIPS:YER043c
 A:Cross-references: SGD:S0000845; MIPS:YER043c
 A:Map position: 5R
 C:Superfamily: adenosylhomocysteinase
 C:Keywords: NAD; thioether hydrolase
 F:214-243/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:80,114/Active site: Cys #status predicted

Query Match 71.9%; Score 1633; DB 2; Length 449;
 Best Local Similarity 70.8%; Pred. No. 1.2e-112;
 Matches 313; Conservative 52; Mismatches 61; Indels 16; Gaps 3;

QY 7 YKVDIGLAAGKGRKALDIAENEMPGLMRERYASAKPLKGARIAGCLHMTTQTAVL 66
 Db 8 YKADIADSLAAGKGRKALDIAENEMPGLMRERYASAKPLKGARIAGCLHMTTQTAVL 67
 QY 67 LVTLGAQVWSSCNIFSTQDQAAAAAIAAGIPVYAWKGTDEEYLWCIEQTLVFKDG-P 124
 Db 68 LVALGAQVWSSCNIFSTQDQAAAAAIAAGIPVYAWKGTDEEYLWCIEQTLVFKDNNK 127
 QY 125 LNMILDDGGDLTNLHITKYPQLPGIRGISETTTGVHNLKYMANGILKVPAINVNDV 184
 Db 128 LNLILDDGGDLTNLHITKYPQLPGIRGISETTTGVHNLKYMANGILKVPAINVNDV 187
 QY 185 VTKSKFDNLGCRSLIDGIKRATDVMIAKGVAVAGYGDVGKCAQALRGFGARVITE 244
 Db 188 VTKSKFDNLGCRSLIDGIKRATDVMIAKGVAVAGYGDVGKCAQALRGFGARVITE 247
 QY 245 DPINALQAMGEGYVTTWDEACQEGNIFVTTTGCIDIIILGRHFEOMKDDAIVCNIGHFD 304
 Db 248 DPINALQAMGEGYVTTWDEACQEGNIFVTTTGCIDIIILGRHFEOMKDDAIVCNIGHFD 307
 QY 305 EIDVKWLNENAVEKNIKPQVDYRLKNGRRITLLAEGRLVNLGCGAMGHPSPVMSNSFTN 364
 Db 308 EIDVAVLNLANA-KKQTVKPVQVDYRLKNGRRITLLAEGRLVNLGCGATGHPSPVMSNSFC 367
 QY 365 QVMAQIELWTHPDYKVPVGVHFLPKKLDDEAVAEHLGKLNVLTKLTKTEKQAOYLGMSCDG 410
 Db 368 QVLAQIALMTWTESEYPLGVHLLPKKLDDEAVARHLGKLVKLTLLTKTEKQSEYLSVPVAG 427
 QY 411 EKQAOYLGMSCDGPFPKPDHYRY 432
 Db 428 KVQSEYLSVPVAGFPKPDHYRY 449

RESULT 8
 A45569

A;Cross-references: GB:AE003941; GB:AE003849; NID:g9105978; PIDN:AAF93847.1; GSPDB:GN00011
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, P.A.; Acencio, M.; Alvarenga, R.; Al
Briiones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorry, H.; Pacincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Macsekuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaasak
A;Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1037
C;Superfamily: adenosylhomocysteinase

Query Match 59.1%; Score 1342; DB 2; Length 446;
Best Local Similarity 61.0%; Pred. No. 3.2e-91;
Matches 272; Conservative 49; Mismatches 83; Indels 42; Gaps 6;

Qy 29 MPGLMRVRYASKPKLGARIACLHMTVETAVLIETVLTLGABVQWSSCNIPSTQNH 88
Db 1 MPGLMSRRKYASKQPLKGVRTVGLSHMTIQTAVLIIETLKDIGADVRWASCNIPSTQDHA 60
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaasak
A;Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1037
C;Superfamily: adenosylhomocysteinase

Qy 89 AAAIAKAGIPVYAWKGTDEYLWCIEQTLYF--KDGPL---NMILDGGDLTNLIHTKY 143
Db 61 AAAIATSGTVPFAWKGETLEBYWDCITLQALTFTLSDGLTGPPELIVDDGGDATTLIHKGY 120
Qy 144 -----PQLL-----PGI-----RGISEETTTGVHNLKYM 167
Db 121 ELENGSTWVDEPSPDLSEQVTKRLKRIATERPQWTVRVNDKGVSEETTTGVHRLVQI 180
Qy 168 MANGILKVPAINVNDSTVKFNDLYGCRSLIDIGIKRATDVMITAGKAVVAVAGYGDVGKG 227
Db 181 AATGRLLVPAINVNDSTVKFNDLYGCRSLADGLKRAMDVMLAGKLAVVCYGDVGKG 240
Qy 228 CAQALRGFGARVITETDPINALQAAMEGYVTTMDACQEGNIFVTTTGCIDITLGRHF 287
Db 241 SAHSLRAYGARVITETDIPCALQAAMEGFVTTVEDTLGQADIYVTTTGNKDVIRIEHM 300
Qy 288 EQMKDDAIVCNIGHFDVEIDVKNLNA--VEKVNIPQVDVRLKNGRRRIILLASRLVN 346
Db 301 TAMKQDVIVCNIGHFDNEIQDALNALAGVQKINIKPQVDKPIPLNGWTLFLASRLVN 360
Qy 347 LGCAMGHPSFVMSNFTNQVMAQIELWTHPDQYPVGVHFLPKCLDEAVAEAHGLKLNVL 406
Db 361 LGCATGHPSFVMSNFTNQVMAQIELWTHPDQYPVGVHFLPKCLDEAVAEAHGLKLNVL 420
Qy 407 TKLTEKQAQYLGMSCDGFPKPDHYRY 432
Db 421 TLTANQAAYLGISVEGFPKPEHYRY 446

RESULT 10
S38379
adenosylhomocysteinase (EC 3.3.1.1) - Madagascan periwinkle
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Catharanthus roseus (Madagascan periwinkle)
C;Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C;Accession: S38379
R;Schroeder, G.; Waitz, A.; Hotze, M.; Schroeder, J.
submitted to the EMBL Data Library, October 1993
A;Description: cDNA for stress-induced S-adenosyl-L-homocysteine hydrolase from Madagasc
A;Reference number: S38379
A;Accession: S38379
A;Molecule type: DNA
A;Residues: 1-485 [SCH]
A;Cross-references: EMBL:Z26881; NID:g407411; PIDN:CAA81527.1; PID:g407412
A;Genetics:

A;Gene: SAHH
 C;Superfamily: adenosylhomocysteinase
 C;Keywords: NAD; thioether hydrolase
 F;263-292/Region: beta-alpha-beta NAD nucleotide-binding fold
 F;86,120/Active site: Cys #status predicted

Query Match 58.3%; Score 1323; DB 2; Length 485;
 Best Local Similarity 57.6%; Pred. No. 9.2e-90;
 Matches 272; Conservative 52; Mismatches 102; Indels 46; Gaps 5;

QY 7 YKVDIGLAAGKALDIAENEMPGMRMRYSASKPLKGARIAGLHMTVETAVLIET 66
 DB 14 YKVKDMSQADFGLEIEAEEVEMPGMLSCRAEFGSPQPFKAKITGSLHMTIQTAVLIET 73
 QY 67 LVTLGAEVQWSSCNIFSTQNHAAATAKAGIPYAWKGETDEEYLWCIEQTL-YFKDGPL 125
 DB 74 LTALGAEVWCSNIFSTQEHAAAIARDAAVFAWKGETLOEYWWCTERALDWDGPDGP 133
 QY 126 NMILDDGGDLTNIH-----TKYP 144
 DB 134 DLIVDDGGDATLIIHGVKAEYKKGALPDSSTDNAEFQIVLTIIRDGLKSDPTKT 193
 QY 145 QLLPGIRGISEETTTGVHNLKYMANGILKVPAINVNDVSKFNLGYCRESLIDGK 204
 DB 194 RMKERLVGVSEETTTGVKELYQMQANGTLFPAINVNDVSKFNLGYCRESLIDGK 253
 QY 205 RATDVMIAKVAAGVGVGKCAALRGFGARVITTEIDPINALQAMGVEYVTMDE 264
 DB 254 RATDVMIAKVAAGVGVGKCAALQKAGARVITTEIDPICALQATMEGLQVLTLED 313
 QY 265 ACEGNIFFVTTCIDIIILGRHPEOMKDAIVCNIGHFVDEIDVKNLE-NAVEKVNIRP 323
 DB 314 VYSEADIFVTTCIDIIIVDHRKMKNAIVCNIGHFNEIDMLGLEYPGVKRITIRP 373
 QY 324 QVDVRYLKN-GRRRIILAEGLVNLGCAMGHPFSFVMSNSTQVMAQIELWTH--PDKYP 380
 DB 374 QTDVWVFPDTSNGIILAEGLVNLGCATGHPFSFVMSNSTQVMAQIELWTH--PDKYP 433
 QY 381 VGVHFLPKLDEAVAEHLKLVNKLTKLTKAQAYLGMSCDGPFPKPDHYRY 432
 DB 434 KKYYVLPKHLDEKVAALHLGKGLAKLTKLTKAQAYLGMSCDGPFPKPDHYRY 485

RESULT 11
 B70593
 adenosylhomocysteinase (EC 3.3.1.1) - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C;Accession: B70593
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: B70593
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-495 <COL>
 A;Cross-references: GB:Z95121; GB:AL123456; NID:G3261742; PIDN:CAB08349.1; PID:G2072715
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: sahH
 C;Superfamily: adenosylhomocysteinase
 C;Keywords: NAD; thioether hydrolase
 F;276-305/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 57.9%; Score 1315.5; DB 2; Length 495;
 Best Local Similarity 55.8%; Pred. No. 3.4e-89;
 Matches 266; Conservative 61; Mismatches 101; Indels 49; Gaps 5;

QY 5 LPYKVADIGLAAGKALDIAENEMPGMRMRYSASKPLKGARIAGLHMTVETAVLI 64

DB 19 IDFKIADLSLADFGKRLRIAEHEMPEGLMSLRREYAEVQPLKGRISGSLHMTVQTAVLI 78
 QY 65 ETLVTLGAEVQWSSCNIFSTQNHAAATIA-----KAGIPYAWKGETDEEYLWCIE 115
 DB 79 ETLTALGAEVWCSNIFSTQDHAAAAVVVVGPHGTPDEPKGVFVFAWKGETLLEEYWAAE 138
 QY 116 QTLYF--KDGPLNMIILDDGGDLT-----NLHIT-- 141
 DB 139 QMLTWPDPDPKPNMILDDGGDATMLVLRMQYKAGVVPAAEDDDPAEWKVFNLRLTRFP 198
 QY 142 -----KYPOLLPGIRGISEETTTGVHNLKYMANGILKVPAINVNDVSKFNLGYCGR 196
 DB 199 ETDKDKWTKIAESVKGVTETTTGVRLYQFAAGDLAFLPAINVDSVTSKFDNKGTR 258
 QY 197 ESLIDGIRKATDVMIAKVAAGVGVGKCAALRGFGARVITTEIDPINALQAMG 256
 DB 259 HSLIDGINRGTDALIGCKKVLICGYGKCAEAMKQCARVSVTEIDPINALQAMG 318
 QY 257 YEVTTMDEACQSGNIFVTTCIDIIILGRHPEOMKDDAIVCNIGHFVDEIDVKNLENAV 316
 DB 319 FQVTVVEEAIGADIVVTATGNKDIIMLEHIKAMKDHAILGNIGHFNDIAGLERSGA 378
 QY 317 EKVNIKPQVDVRYL-KNGRRIILAEGLVNLGCAMGHPFSFVMSNSTQVMAQIELWTH 375
 DB 379 TRVNVKQVDLWTFGDTGTSIIIVLSEGLLNGLNATGHPFSFVMSNSTQVMAQIELWTH 438
 QY 376 PKYPVGVHFLPKLDEAVAEHLKLVNKLTKLTKAQAYLGMSCDGPFPKPDHYRY 432
 DB 439 NDEYDNEVYRLPKHLEKVARIHVEALGHLTKLTKAQAYLGMSCDGPFPKPDHYRY 495

RESULT 12
 A46035
 adenosylhomocysteinase (EC 3.3.1.1) - Rhodobacter capsulatus
 N;Alternate names: S-adenosyl-L-homocysteine hydrolase
 C;Species: Rhodobacter capsulatus
 C;Date: 04-Mar-1994 #sequence_revision 04-Oct-1996 #text_change 18-Jun-1999
 C;Accession: A46035; A36863
 R;Sganga, M.W.; Aksamat, R.R.; Cantoni, G.L.; Bauer, C.E.
 Proc. Natl. Acad. Sci. U.S.A. 89, 6328-6332, 1992
 A;Title: Mutational and nucleotide sequence analysis of S-adenosyl-L-homocysteine hydrolase
 A;Reference number: A46035; MUID:92335291; PMID:1631127
 A;Accession: A46035
 A;Molecule type: DNA
 A;Residues: 1-462 <SGA>
 A;Note: sequence extracted from NCBI backbone (NCBIN:108691, NCBIPI:108696)
 R;Buggy, J.J.; Sganga, M.W.; Bauer, C.E.
 J. Bacteriol. 176, 61-69, 1994
 A;Title: Nucleotide sequence and characterization of the Rhodobacter capsulatus hvrB gene
 A;Reference number: A36863; MUID:94110241; PMID:8282711
 A;Accession: A36863
 A;Molecule type: DNA
 A;Residues: 1-113 <BUG>
 A;Cross-references: GB:L23836; NID:G577613; PIDN:AAA53540.1; PID:G577614
 C;Genetics:
 A;Gene: hcy
 C;Function:
 A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to adenosine
 C;Superfamily: adenosylhomocysteinase
 C;Keywords: NAD; thioether hydrolase
 F;247-276/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 57.8%; Score 1312; DB 1; Length 462;
 Best Local Similarity 58.0%; Pred. No. 5.5e-89;
 Matches 275; Conservative 49; Mismatches 96; Indels 54; Gaps 6;

QY 1 MSDKLPYKVADIGLAAGKALDIAENEMPGMRMRYSASKPLKGARIAGLHMTVET 60
 DB 1 MAD-----YIVKDKIAEFGKELDIAETEMPGLMACREFGSPQPLKGRISGSLHMTIQT 57
 QY 61 AVLIETLTLGAEVQWSSCNIFSTQNHAAATAKAGIPYAWKGETDEEYLWCIEQTL 120

Db 58 AVLIELTKALGADVRWASCNIFSTQDHAIAAIAAGTTPVFAVKGETLBEYWAYTKIFQF 117
 QY 121 KQGLPNMILDDGGDL-----TNLIHTK-----XP 144
 Db 118 PEGTCNMILDDGGDATLYILLGARVENGETDLIATPSEDEVCLFNQIKRWVESGWT 177
 QY 145 QLLPGIRGISEETTTGVHNLKYMANGILKVPAINVNDSTKSFNLYGCRSLIDGK 204
 Db 178 QQRAAIKGVSEETTTGVHRLYDLKKGKLLPFPAINVNDSTKSFNLYGCRSLVDGIR 237
 QY 205 RATDVMIAKVAVAGYGVGKCAQALRGFGARVITTEIDPINALQAAAMEGVVTTMDE 264
 Db 238 RATDVMIAKVAVAGYGVGKCAQALRGFGARVITTEIDPINALQAAAMEGVVTTMDE 297
 QY 265 ACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHFDVEIDV-----KMLNENAVEK 318
 Db 298 VVADA-IFITTTGNKDVIREHWRMKDMAIVGNIGHFDNEIOVAALKNHKW-----348
 QY 319 VNIKPQVDYRLKNGRRILLAEGRVLNLCGAMGHPSPVMSNFTNQVMAQIELWTHPK 378
 Db 349 TNIKQDQMIEMPSGARIILLSEGRLLNLGNATGHPSPVMSNFTNQVMAQIELWTKGAE 408
 QY 379 YPVGVHFLPKLDEAVAEHLGKLVNLTETKQAOYLGMSCDGPFPKPDHYRY 432
 Db 409 YQGVVILPKSLDEKVARLHLKKGKLVTLRDPQAEYIGVTVEGPFKSDHYRY 462
 RESULT 13
 D87005
 probable S-adenosyl-L-homocysteine hydrolase [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: D87005
 R: Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H
 R.; Davies, R.M.; Devlin, K.; Duthey, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: A86909; MUID:21128732; PMID:11234002
 A:Accession: D87005
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-492 <STO>
 A:Cross-references: GB:AL450380; NID:G13092881; PIDN:CAC30280.1; GSPDB:GN00147
 A:Gene: sahH
 C:Superfamily: adenosylhomocysteinease
 Query Match 57.4%; Score 1302.5; DB 2; Length 492;
 Best Local Similarity 55.6%; Pred. No. 3e-88;
 Matches 265; Conservative 58; Mismatches 105; Indels 49; Gaps 4;
 QY 5 LPYKVADIGLAAGWRKALDAENMPGLMRERYASAKPLKGARIAGCLHMTVETAVLI 64
 Db 16 IDPKVADLSANFGKRLDLAEYMPGLMSLRHEYAEVQPLKGARISGLHMTVQTAVALI 75
 QY 65 ETLLTALGAEVQWSSCNIFSTQDHAIAAIA-----KAGIPVYAWKGETDEBYLWCI 115
 Db 76 ETLLTALGAEVQWSSCNIFSTQDHAIAAIAVVGYPGTPPEPKGVFPFAWKGETLEBYWAAE 135
 QY 116 QTLYF--KQGLPNMILDDGGDLTNLI-----139
 Db 136 QMLTWFPDPKPVNMILDDGGDATMLVLRGVQYKAGVVPFAEVDDSAEWKVFNLRLKRP 195
 QY 140 ---HTKYPQLLPGIRGISEETTTGVHNLKYMANGILKVPAINVNDSTKSFNLYGCR 196
 Db 196 ETDKQWTKIAKSVKGVTEETTTGVRLYQFAAAGDLAPFPAINVNDSTKSFNLYGCR 255
 QY 197 ESLIDGIRKATDVMIAKVAVAGYGVGKCAQALRGFGARVITTEIDPINALQAAAMEG 256
 Db 256 HSLIDGINRGTDLSLIGKKNVLCGCGYGVGKCAEAAGQAGQARTITEIDPINALQALMEG 315

QY 257 YEVTMTDEACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHFDVEIDVYKMLNENAV 316
 Db 316 PDVKRVEDVIADSDIVVTATGNKDIILLEHMKAMKDHAILGNIGHFDNEIDMAALERSGA 375
 QY 317 EKVNIPQVDYRL-KNGRRILLAEGRVLNLCGAMGHPSPVMSNFTNQVMAQIELWTH 375
 Db 376 TRLNIPQVDLWTFDGSKSIIVLSEGRLLNLGNATGHPSPVMSNFTNQVMAQIELWTK 435
 QY 376 PDKYVGVHFLPKLDEAVAEHLGKLVNLTETKQAOYLGMSCDGPFPKPDHYRY 432
 Db 436 NDDYDNEVYRLPKHLDEKVARVHVZALGGQLTKLTKDQAEYLVGDVDPKPDHYRY 492
 RESULT 14
 AG3505
 adenosylhomocysteinease (EC 3.3.1.1) [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
 R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mijar, C.; Los, T.; Ivanova, I
 R.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesec
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AG3505; PMID:11756688
 A:Accession: AG3505
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-481 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AAL53210.1; PID:G17984086; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Map position: 1
 C:Superfamily: adenosylhomocysteinease
 C:Keywords: thioether hydrolase
 Query Match 57.3%; Score 1302; DB 2; Length 481;
 Best Local Similarity 57.2%; Pred. No. 3.2e-88;
 Matches 271; Conservative 52; Mismatches 101; Indels 50; Gaps 6;
 QY 1 MSDKLPYKVADIGLAAGWRKALDAENMPGLMRERYASAKPLKGARIAGCLHMTVET 60
 Db 16 MTASQDFVVKDIISLADWGRRELDIAETEMPGMLAAREEFKSGPLKARISGLHMTIQT 75
 QY 61 AVLIELTALGAEVQWSSCNIFSTQDHAIAAIA-----KAGIPVYAWKGETDEBYLWCI 120
 Db 76 AVLIELTKVLGAEVQWSSCNIFSTQDHAIAAIAATGTPVFAVKGETLEBYTWTYDQIFOW 135
 QY 121 KQGLPNMILDDGGDLTNLI-----HTKYPQL-----LPG---149
 Db 136 PDGEPNMLDDGGDATMYILLGARAEAGEDVLSNPQSEEEVLPFAIKKRWMAATGPTT 195
 QY 150 -----IRGISEETTTGVHNLKYMANGILKVPAINVNDSTKSFNLYGCRSLIDGK 204
 Db 196 KQRAAIKGVTEETTTGVHNLKYMANGILKVPAINVNDSTKSFNLYGCRSLIDGK 255
 QY 205 RATDVMIAKVAVAGYGVGKCAQALRGFGARVITTEIDPINALQAAAMEGVVTTMDE 264
 Db 256 RGTDMVMAKVAVAGYGVGKCAQALRGFGARVITTEIDPINALQAAAMEGVVTTMDE 315
 QY 265 ACQEGNIFVTTTGCIDIIILGRHFEQMKDDAIVCNIGHFDVEIDV-----KMLNENAVEK 318
 Db 316 AASTADIVTTTGNKDVIREHWRMKDMAIVGNIGHFDNEIOVAALKNHKW-----367
 QY 319 VNIKPQVDYRLKNGRRILLAEGRVLNLCGAMGHPSPVMSNFTNQVMAQIELWTHPK 378
 Db 368 TNVKEQVDLIEPDKRLILLSEGRLLNLGNATGHPSPVMSNFTNQVMAQIELWTRDA 427
 QY 379 YPVGVHFLPKLDEAVAEHLGKLVNLTETKQAOYLGMSCDGPFPKPDHYRY 432
 Db 428 YKNEVYVLPKHLDEKVARLHLKKGKLVTLRDPQAEYIGVTVEGPFKSDHYRY 481

RESULT 15

AF2580
S-adenosylhomocysteine hydrolase ahcy [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AF2580
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, B.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:21608550; PMID:11743193
A:Accession: AF2580
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-466 <KUR>
A:Cross-references: GB:AE008688; PIDN:AA141060.1; PID:gl7738347; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: ahcy
A:Map position: circular chromosome
C:Superfamily: adenosylhomocysteinase

	Query Match	57.2%	Score 1299;	DB 2;	Length 466;
	Best Local Similarity	57.6%;	Pred. No. 5.1e-88;		
	Matches 273;	Conservative 51;	Mismatches 100;	Indels 50;	Gaps 5;
Qy	1	MSDKLPYKVDIGLAAMGRKALDIAENEMPGLMRMERYASKPLKAGARIAGCLHMTVET	60		
Db	1	MSLEKDYIVADINLAAGFKELDIAETEMEGLMSCRFESKPLKAGARIAGCLHMTVET	60		
Qy	61	AVLIETLVILGAEVQSSCNIFSTQNHAAAAIAKAGIPYAMKGETDEBYLMCIQOTLYP	120		
Db	61	GVLIETLKEIGAIRWASCNIFSTQDHAAAAIAAGIPVFAVKGESLTYEYWEYTKIFQW	120		
Qy	121	KDGPL-NMILDDGGDLTNLI-----HTKYPQLLPG----	149		
Db	121	TGGLSNMILDDGGDAITMVLIGARAEAGEDVLSNPGSEBEILPAQINKRLKASPGWFT	180		
Qy	150	-----IRGISEETTGVHNLKMMANGILKVPAINVNDSVTKSKPDNLYGCRESLIDGK	204		
Db	181	KORDALKGVTETTTGVHRLYDLAKGGLPPFPAINVNDSVTKSKPDNLYGCRESLVDGIR	240		
Qy	205	RATDVMIAKVAVAGYGVGKCAQALRGFGARVITIDIPINALQAAMEGVEYVTMD	264		
Db	241	RATDVMAGKVAVCGYGVGKSAASLQAGARVKVTEIDPICALQAAMDGFVVRL	300		
Qy	265	ACQEGNIFVTTCIDILGRHFEQMKDDAIVCNIGHFDVEIDV-----KMLNENAVEK	318		
Db	301	VVSSADIFITTTGNKDVIRIEHREMKDMAIVGNIGHFDNEIQVASLRNLK-----	352		
Qy	319	VNIKPVDRYRLKNGRRIILAEGRVLNLGCAMGHPSPVMSNFTNQVMAQIELWTHPK	378		
Db	353	TNIKPVDMIEPPKGRRIILSSEGRLLNLGNATGHPSPVMSASFTNQVLQIELFTKPC	412		
Qy	379	YPVGVHFLPKLDEAVAEAHGLKLVKLTETKQAQYLGMSCDGPFKPDHYR	432		
Db	413	YKNEVYLPKHLDEKVARLHLEKLVRLTELSLDQADYIGISKQGFKAHYR	466		

Search completed: December 30, 2003, 10:42:32
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2003, 10:36:48 ; Search time 46 Seconds
(without alignments)
1490.650 Million cell updates/sec

Title: US-10-043-787-1

Perfect score: 2271

Sequence: 1 MSDKLPYKVDIGLAAGRK.....QAQLGMSCDGPFKPDHYRY 432

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2266	99.8	457	21	AA43614 Human cancer assoc
2	2204.5	97.1	433	23	AA25380 Human NZMS-4 prote
3	1819.5	80.1	432	22	AB70895 Drosophila melanog
4	1393	61.3	469	24	ABU11410 Protein encoded by
5	1315.5	57.9	495	21	AAV87873 M. tuberculosis an
6	1297	57.1	485	21	AA34060 Zea mays protein f
7	1297	57.1	510	21	AA34059 Zea mays protein f
8	1296	57.1	486	13	AA26500 Prod. of Nicotiana
9	1292	56.9	485	21	AA40086 Arabidopsis thalia

10	1292	56.9	508	21	AA40085 Arabidopsis thalia
11	1289	56.8	485	17	AAW01458 Aspartag S-adenos
12	1274	56.1	485	21	AA41407 Arabidopsis thalia
13	1274	56.1	497	21	AA41406 Arabidopsis thalia
14	1272	56.0	467	21	AA40087 Arabidopsis thalia
15	1254	55.2	467	21	AA41408 Arabidopsis thalia
16	1248	55.0	450	21	AAU34061 Zea mays protein f
17	1242	54.7	478	22	AAU71911 C. glutamicum meta
18	1242	54.7	478	22	AAU71911 C. glutamicum meta
19	1242	54.7	498	23	AAU78423 Corynebacterium g1
20	1238	54.5	474	22	AAU78423 Corynebacterium g1
21	1236	54.4	485	17	AAW01459 Arabidopsis S-aden
22	1224	53.9	467	21	AAW01459 Arabidopsis thalia
23	1224	53.9	489	21	AAW01459 Arabidopsis thalia
24	1204	53.0	449	21	AAW01459 Arabidopsis thalia
25	1151	50.7	614	19	AAW56097 Amino acid sequenc
26	1149.5	50.6	504	22	AB71710 Drosophila melanog
27	1149.5	50.6	521	22	AB71710 Drosophila melanog
28	1146	50.5	500	20	AAW90061 Human SAHH protein
29	1139	50.2	564	23	ABG32371 Novel human enzyme
30	1098	48.3	432	22	AAU71913 C. glutamicum meta
31	1098	48.3	432	22	AAU71913 C. glutamicum meta
32	1084	47.7	395	23	AAE23390 Human NZMS-14 prot
33	878.5	38.7	422	22	AAW98504 Pucative P. abyssi
34	870	38.3	170	23	AAW98504 Rat S-adenosylhomo
35	779	34.3	455	22	ABG12029 Novel human diageno
36	761	33.5	464	21	AAU77950 A. thaliana enviro
37	423	18.6	500	23	ABP65910 Bifidobacterium lo
38	416.5	18.3	410	22	ABG27861 Novel human diageno
39	415	18.3	81	21	AAW01918 Human secreted pro
40	415	18.3	97	16	AAW01918 Human SAHH. Homo
41	394	17.3	95	16	AAW01918 Human secreted p
42	385	17.0	135	22	AAO10323 SAHH-GSE-encoded p
43	296	13.0	66	23	ABP42348 Human polypeptide
44	291	12.8	89	23	ABP42348 Human ovarian anti
45	276.5	12.2	312	22	ABG12030 Novel human diageno

ALIGNMENTS

RESULT 1

AA43614

ID AA43614 standard; Protein; 457 AA.

XX AA43614;

AC AA43614;

XX AA43614;

DT 08-FEB-2001 (first entry)

XX Human cancer associated protein sequence SEQ ID NO:1059.

DE Human; cancer associated gene; cancer antigen; detection; cancer;

XX Human; cancer associated gene; cancer antigen; detection; cancer;

KW diagnosis; cyclostatic; proliferative; vulnerable; immunomodulator;

KW antidiabetic; antiaesthetic; antithrombotic; antithrombotic; antiviral;

KW antinflammatory; antithyroid; antiallergic; antibacterial; cardiac;

KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;

KW vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;

KW immune disorder; haematopoietic cell disorder; autoimmune disorder;

KW allergic reaction; graft versus host disease; organ rejection;

KW haemostatic; thrombolytic; cardiovascular disorder; infection;

KW neurologic disease; drug screening.

XX Homo sapiens.

OS Homo sapiens.

XX WO20005350-A1.

FN WO20005350-A1.

XX 21-SEP-2000.

PD 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05882.

PF 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

PR 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;
 PI WPI; 2000-587533/55.
 DR N-PSDB; AAC77823.
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 XX Claim 11; Page 1650-1652; 2352pp; English.
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnary; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX Sequence 457 AA;

Query Match 99.8%; Score 2266; DB 21; Length 457;
 Best Local Similarity 99.8%; Pred. No. 5e-226;
 Matches 431; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDKLPYKVDIGLAAGRWKALDIAENEMPLMRERYASPLKRGARIAGLHMTVET 60
 DB 26 MSDKLPYKVDIGLAAGRWKALDIAENEMPLMRERYASPLKRGARIAGLHMTVET 85
 QY 61 AVLIETLVLTGAEVQSSCNIFSTQNHAAIAKAGIPYVANKGETDEEVLWCIEQTLYF 120
 DB 86 AVLIETLVLTGAEVQSSCNIFSTQNHAAIAKAGIPYVANKGETDEEVLWCIEQTLYF 145
 QY 121 KDGPLNMILDGDLTNLIHTKYPQLPGIRGISEETTTGVHNLKXWANGILKVPAINV 180
 DB 146 KDGPLNMILDGDLTNLIHTKYPQLPGIRGISEETTTGVHNLKXWANGILKVPAINV 205
 QY 181 NDSVTKSPNLYGCRESLIDGIKRATDVMIAGKAVVAGYGVGKGCQAALGFQARVI 240
 DB 206 NDSVTKSPNLYGCRESLIDGIKRATDVMIAGKAVVAGYGVGKGCQAALGFQARVI 265
 QY 241 ITEIDPINALQAAMEGVETVMTDEACQEGNIFVTTTGCIDIIILGRHFQMKDAIVCNIG 300
 DB 266 ITEIDPINALQAAMEGVETVMTDEACQEGNIFVTTTGCIDIIILGRHFQMKDAIVCNIG 325
 QY 301 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGRVLNLCAMGHPSPVMSN 360
 DB 326 HFDVEIDVKWLNENAVEKVNIPQVDYRLKNGRRILLAEGRVLNLCAMGHPSPVMSN 385
 QY 361 SFTNQWMAQIELTHDPKDPVGVHFLPKLDEAVAHAHGLKLVKLTKEQAQYLGMS 420
 DB 386 SFTNQWMAQIELTHDPKDPVGVHFLPKLDEAVAHAHGLKLVKLTKEQAQYLGMS 445
 QY 421 CDGPFKPDHYRY 432
 DB 446 CDGPFKPDHYRY 457

RESULT 2

AAE25380
 ID AAE25380 standard; Protein; 433 AA.
 XX AAE25380;
 AC AAE25380;
 XX 30-OCT-2002 (first entry)
 DT XX
 XX Human NZMS-4 protein.
 DE XX
 XX Human; enzyme; NZMS-4; cell proliferative disorder; hepatitis; cancer;
 KW psoriasis; leukaemia; autoimmune disorder; diabetes; arteriosclerosis;
 KW acquired immune deficiency syndrome; cardiovascular disorder; anorexia;
 KW asthma; hypertension; gastrointestinal disorder; reproductive disorder;
 KW gastritis; neurological disorder; epilepsy; pulmonary disorder; AIDS;
 KW dementia; embolism; gene therapy; eye disorder; transgenic.
 XX Homo sapiens.
 OS WO200246385-A2.
 PN 13-JUN-2002.
 PD XX
 XX 04-DEC-2001; 2001WO-US47432.
 PF XX
 XX 07-DEC-2000; 2000US-251824P.
 PR 08-DEC-2000; 2000US-254312P.
 PR 14-DEC-2000; 2000US-255773P.
 PR 15-DEC-2000; 2000US-255940P.
 PR 15-DEC-2000; 2000US-256188P.
 PR 21-DEC-2000; 2000US-257488P.
 PR 19-JAN-2001; 2001US-262839P.
 PR 26-JAN-2001; 2001US-264402P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Tang YT, Griffin JA, Yue H, Lee EA, Baughn MR, Duggan BM;
 PI Walla NK, Lee S, Ramkumar J, Warren BA, Gandhi AR, Lu Y;
 PI Yao MG, Ding L, Tribouley CM, Sanjanwala MM, Arvizu C, Hillman JL;
 XX WPI; 2002-537565/57.
 DR N-PSDB; AAB411391.
 XX Novel human enzyme, NZMS useful in diagnosis, prevention or treatment
 PT of cell proliferative, autoimmune/inflammatory, cardiovascular,
 PT gastrointestinal, neurological, pulmonary, reproductive and eye
 PT disorders -
 XX Claim 59; Page 139-140; 173pp; English.
 PS The invention relates to human enzymes designated NZMS and nucleic
 CC acid molecules encoding such proteins. Sequences of the invention
 CC are useful for diagnosing, treating or preventing disorders associated
 CC with aberrant expression of NZMS. The disorders treated include
 CC cell proliferative disorders such as hepatitis, psoriasis, cancer
 CC (e.g. leukaemia), autoimmune disorders such as diabetes, acquired
 CC immune deficiency syndrome (AIDS), cardiovascular disorders such as
 CC arteriosclerosis, hypertension), gastrointestinal disorders (e.g.
 CC anorexia, gastritis), neurological disorders (e.g. epilepsy, dementia),
 CC pulmonary disorders (e.g. embolism, asthma), reproductive or eye
 CC disorders. Polypeptides of the invention is useful in a number of drug
 CC screening techniques and to analyse the proteome of a tissue or cell
 CC type. They are also useful as elements on a microarray. Polynucleotides
 CC of the invention are useful for creating knockin humanised animals or
 CC transgenic animals to model human diseases, in somatic or germline gene
 CC therapy, to generate a transcript image of a tissue or cell type, for
 CC detecting difference in the chromosomal location due to translocation
 CC or inversion among normal, carrier or affected individuals and as
 CC hybridisation probes for mapping naturally occurring genomic sequences.
 CC The present sequence is human NZMS protein.
 XX Sequence 433 AA;

Query Match 97.1%; Score 2204.5; DB 23; Length 433;

Best Local Similarity 97.2%; Pred. No. 1.1e-219; Matches 421; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 MSKLPYKADIGLAAGKALDIAENEMPGLMRERYASAKPLKARIAGLHMTVET 60
DB 1 MSKLPYKADIGLAAGKALDIAENEMPGLMRERYASAKPLKARIAGLHMTVET 60
QY 61 AVLIETLVITGAQVOWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIEQTLVF 120
DB 61 AVLIETLVITGAQVOWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIEQTLVF 120
QY 121 KOGPLNMLDDGGDLNLHTKYPQLLPGIRGISEETTTGVHNLVYOMANGILKVPAINV 180
DB 121 KOGPLNMLDDGGDLNLHTKYPQLLPGIRGISEETTTGVHNLVYOMANGILKVPAINV 180
QY 181 NDSVTK-SKFDNLVYGCRESLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARV 239
DB 181 NDSVTK-SKFDNLVYGCRESLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARV 240
QY 240 IITEIDIPNALQAAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFQMKDDAIVCNI 299
DB 241 IITEIDIPNALQAAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFQMKDDAIVCNI 300
QY 300 GHFDVIDVKWLNENAVEKNIKPQVDYRLKNGRRIILAEGLRVNLGCAMGHPSPFVMS 359
DB 301 GHFDVIDVKWLNENAVEKNIKPQVDYRLKNGRRIILAEGLRVNLGCAMGHPSPFVMS 360
QY 360 NSFTNQVMAQIELWTHPDKYPGVHPLPKLDVAEVAHLGKLVNKLTKLTKAQOYLGM 419
DB 361 NSFTNQVMAQIELWTHPDKYPGVHPLPKLDVAEVAHLGKLVNKLTKLTKAQOYLGM 420
QY 420 SCGPPKPDHRY 432
DB 421 PIDGPPKPDHRY 433

RESULT 3
ABB70895
ID ABB70895 standard; Protein; 432 AA.
AC ABB70895;
XX
XX
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 39477.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL14998.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 39477; 21pp + Sequence Listing; English.
XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 432 AA;

Query Match 80.1%; Score 1819.5; DB 22; Length 432;
Best Local Similarity 80.7%; Pred. No. 1.1e-179; Indels 1; Gaps 1;
Matches 347; Conservative 28; Mismatches 54;

QY 4 KLPYKADIGLAAGKALDIAENEMPGLMRERYASAKPLKARIAGLHMTVETAVL 63
DB 3 KPSYKADISLAEGWKALIIIAENEMPGLMACRKKYPSKPLKARITGCLHMTVQTAVL 62
QY 64 IETLVITGAQVOWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIEQTLVFKDG 123
DB 63 IETLVITGAQVOWSSCNIFSTQNHAAAIKAGIPVYAWKGETDEEYLWCIEQTLVFPDG 122
QY 124 -PLNMLDDGGDLNLHTKYPQLLPGIRGISEETTTGVHNLVYOMANGILKVPAINVND 182
DB 123 QPLNMLDDGGDLNLVHEKFPQYLNKIKGLSEETTTGVHNLVYOMKPEGRGLGVPAINVND 182
QY 183 SVTKSKFDNLVYGCRESLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVIT 242
DB 183 SVTKSKFDNLVYGCRESLIDGIKRATDVMIAGKAVVAGYGVGKCAQALRGFGARVIT 242
QY 243 EIDPINALQAAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFQMKDDAIVCNIH 302
DB 243 EVDPINALQAAAMEGYEVTMTDEACQEGNIFVTTTGCIDIIILGRHFQMKDDAIVCNIH 302
QY 303 DVEIDVKWLNENAVEKNIKPQVDYRLKNGRRIILAEGLRVNLGCAMGHPSPFVMSNSF 362
DB 303 DVEIDVKWLNENAVEKNIKPQVDYRLKNGRRIILAEGLRVNLGCAMGHPSPFVMSNSF 362
QY 363 TNQVMAQIELWTHPDKYPGVHPLPKLDVAEVAHLGKLVNKLTKLTKAQOYLGMSCD 422
DB 363 TNQVMAQIELWTHPDKYPGVHPLPKLDVAEVAHLGKLVNKLTKLTKAQOYLGMSCD 422
QY 423 GPPKPDHRY 432
DB 423 GPPKPDHRY 432

RESULT 4
ABU11410
ID ABU11410 standard; Protein; 469 AA.
XX
XX AC ABU11410;
XX
XX
DT 11-FEB-2003 (first entry)
DE
DE
XX Protein encoded by S. atroolivaceus leinamycin gene cluster ORF+8.
XX
XX Leinamycin biosynthesis gene cluster; Lmn; open reading frame; ORF;
KW anti-tumour antibiotic; broad spectrum antimicrobial activity;
KW Gram-positive; Gram-negative bacteria; chemical modification;
KW metabolite; apo-carrier protein; holo-carrier protein; tumour;
KW polyketide; hybrid polyketide/polyketide metabolite; Lmn production;
XX cytosolic.
XX
OS Streptomyces atroolivaceus.
XX
XX WO200277179-A2.
XX
XX 03-OCT-2002.

CC sequence represents a Microbacterium tuberculosis TB54 antigen
described in the invention.

XX
SQ Sequence 495 AA;

Query Match 57.9%; Score 1315.5; DB 21; Length 495;
Best Local Similarity 55.8%; Pred. No. 2.9e-127;
Matches 266; Conservative 61; Mismatches 101; Indels 49; Gaps 5;

QY	5	LPYKADIGLAAGRKALDAENEMQGLMERMERYASASKPLKCARITAGCLHMTVETAVLI	64
DB	19	IDPKIADLSLADGRKELRTAEHEMGLSLREYAEVQPLKGRISGSLHMTVQVAVLI	78
QY	65	ETLVTIGAEVQWSSCNIFSTQNHAAAAIA-----KAGIPVYAWKGETDEEYLWCIE	115
DB	79	ETLTALGAERVWASCNIFSTQDHAAAAVVVPHGTPDEPKGVFVFAWKGETLEEYVWAAE	138
QY	116	QTYLP--KQGLNMLDDGGDLT-----NLHT--	141
DB	139	QMLTWPPDPKPNMILDDGGDATMLVLRGMQYKAGVVPPEEDDPAEWKVFLLRLTRF	198
QY	142	-----KYPQLLPGIRGISERTTGVHNLKWMANGILKVPAINVDSVTKSKEDNLYGCR	196
DB	199	ETDKWTKIAESVKGVTETTTGVRLRYQFAAGDLAPPAINVDSVTKSKEDNLYGTR	258
QY	197	ESLIDGIKRATDVMIAGKAVAGVGDVGKCAQALRGFGARVITTEIDPINALQAAWEG	256
DB	259	HSLLIDGINRGTDALIGKKVLCYGDVGKCAEMKGGQARVSVTEIDPINALQAAWEG	318
QY	257	YEVTWDEACQEGNI FVTTTGCIDIIILGRHFEQKODAIVCNIGHFDEIDVKMLNENAV	316
DB	319	FDDVTVEEAIGDADIVVATGNKDIIMLEHIKAKDHAILGNIGHFDEIDMAGLERSGA	378
QY	317	EKNVIRPOVDYRL-KNGRIILLAEGLVNLGCMGHPSPVMSNFTNQVMAQIELWTH	375
DB	379	TRYNVRPQVDLWTFDGTGRSIIVLSEGLRLNLGNATGHPSPVMSNFPANQIAQIELWTK	438
QY	376	PDKYPGVVHPLPKKLEDAVAEHLKLVNKLTKLTKAQOYLGWSCDGPPEKPDHYRY	432
DB	439	NDEYDNEVYRLPHKLEKVARIHVEALGGHLTKLTKAQOYLGVDVGGPKPDHYRY	495

RESULT 6
AAG34060
ID AAG34060 standard; Protein; 485 AA.
AC AAG34060;
XX
XX
DT 18-OCT-2000 (first entry)
DE Zea mays protein fragment SEQ ID NO: 41387.
XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence; Corn.
XX
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 23-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.

PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	04-MAY-1999;	99US-0132407.
PR	05-MAY-1999;	99US-0132484.
PR	06-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	21-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135153.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.

```
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.1%; Score 1297; DB 21; Length 485;
Best Local Similarity 56.9%; Pred. No. 2,3e-125;
Matches 269; Conservative 52; Mismatches 104; Indels 48; Gaps 6;

QY 7 YKVADIGLAAMGRKALDIAENEMPGLMRMRYRSASKPLKGARIAGCLHMTVETAVLIET 66
DB 14 YKVDLSQADFQRLERLEIAEVEPMGLMACRABFGSKPFAGARISGSLHMTIQTAVLIET 73
QY 67 LVTLGAEVQWSSCNIFSTQNHAAAIKAKGIPVYAWKGTDEYLCWICIBQTIYFKD--GP 124
DB 74 LTALGAEVRCWSSCNIFSTQDHAAAIARDSAVFAWKGTLEEYWWCTERCIDWGEAGGP 133
QY 125 LNMILDDGGDLNLH-----TKY 143
DB 134 -DLIVDDGGDATLLIHGKAEYEYKTKIPDPESTDNABFKIVLTIIRDGLKADPKKY 192
QY 144 POLLGIRIGISETTTGVHNLKMWANGILKVPAINVNDSTKSKFDNLYGCRSLIDGI 203
DB 193 RKMKEKLVGVSETTTGVKRLYQMDETGALLFPAINVNDSTKSKFDNLYGCRSLPDGL 252
QY 204 KBATDVIAGKVVAVVAGYGVKGCAQALRGFGARVITEIDPINALQAAMEGYEYVTMD 263
DB 253 MRATDVIAGKVVAVVAGYGVKGCAALKQAGARVITEIDPICALQALMEGLQVLPLE 312
QY 264 EACQEGNIFVTTTGCIDILGRHFPQMKDDATVCMNIGHFDVIDVKWLN-NAVEKWNK 322
DB 313 DVSSEADIFVTTTGNKDIIMVDHMRKKNNAIVCNIGHFDNEIDMLGLEYTPGVKRITIK 372
QY 323 POVDVRL-KNGRRILLAEGLVNLGCAMGHPSTFVMSNSTNQVMAQIELWTHPD--KY 379
DB 373 PQTDRVFPETWTGIIVLAEGRLMNLGCATGHPSTFVMSNSTNQVIAQELWKESSGKY 432
QY 380 PVGVHFLPKKLDEAAVAEHLGNVLTKLTKTEKQAOYLGMSCDGGFPKPDHYRY 432
DB 433 EKKVVVLPKHLDEKVAALHGLGAKLTKLTKSQADYISVPTEGPKPAHYRY 485

RESULT 7
AAG34059
ID AAG34059 standard; Protein; 510 AA.
XX
AC AAG34059;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 41386.
XX
```



```
OY 205 RATDVMAGKAVVAGVGVKCAQALRGFGARVITEIDPINALQAAMEGYEVTMTDE 264
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 255 RATDVMAGKAVVAGVGVKCAQALRGFGARVITEIDPICALQATMEGLQVLTLED 314
OY 265 ACEGNEIVTTTCIDIILGRHFEQMKDDAIVCNIGHFDVVIDVKWNE-NAVEKYNIXP 323
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 315 VVSDVDFVTTTGNKDIIMVDHMRKMNNAIVCNIGHFDNEIDMLGLEYFGVKKIIIXP 374
OY 324 QVDYRLKN-GRRIIILAEGLVNLGCAMGHPFVMSNSFTNQVMAQIELWTHPD--KYP 380
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 375 QTDWVFPDINSGLIIVLAEGRLNLCATGHPFVMSCSFTNQVIAQLELWNEKSGKYE 434
OY 381 VGVHFLPKLDEAVASAHGLKLVNKTLEKQAQYLGMSCDGPFKPDHYRY 432
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 435 KKYVLPPLHDEKVAALHLGLKGLKAKLTKLSKDQADYISVPVEGYPKPAHYRY 486

RESULT 9
AAG40086
ID AAG40086 standard; Protein; 485 AA.
XX
AC AAG40086;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49691.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
```

PR 04-AUG-1999;	99US-0147302.	Query Match	56.9%;	Score 1292;	DB 21;	Length 485;
PR 05-AUG-1999;	99US-0147192.	Best Local Similarity	56.3%;	Pred. No. 7.7e-125;		
PR 05-AUG-1999;	99US-0147260.	Matches 267;	Conservative 55;	Mismatches 102;	Indels 50;	Gaps 6;
PR 06-AUG-1999;	99US-0147303.					
PR 06-AUG-1999;	99US-0147416.					
PR 09-AUG-1999;	99US-0147493.					
PR 09-AUG-1999;	99US-0147935.					
PR 10-AUG-1999;	99US-0148171.					
PR 11-AUG-1999;	99US-0148319.					
PR 12-AUG-1999;	99US-0148341.					
PR 13-AUG-1999;	99US-0148565.					
PR 13-AUG-1999;	99US-0149229.					
PR 16-AUG-1999;	99US-0149368.					
PR 16-AUG-1999;	99US-0149368.					
PR 17-AUG-1999;	99US-0149175.					
PR 18-AUG-1999;	99US-0149426.					
PR 20-AUG-1999;	99US-0149722.					
PR 20-AUG-1999;	99US-0149723.					
PR 20-AUG-1999;	99US-0149929.					
PR 23-AUG-1999;	99US-0149902.					
PR 23-AUG-1999;	99US-0149930.					
PR 25-AUG-1999;	99US-0150566.					
PR 26-AUG-1999;	99US-0150884.					
PR 27-AUG-1999;	99US-0151065.					
PR 27-AUG-1999;	99US-0151066.					
PR 30-AUG-1999;	99US-0151303.					
PR 31-AUG-1999;	99US-0151438.					
PR 01-SEP-1999;	99US-0151930.					
PR 07-SEP-1999;	99US-0152363.					
PR 10-SEP-1999;	99US-0153070.					
PR 13-SEP-1999;	99US-0153758.					
PR 15-SEP-1999;	99US-0154018.					
PR 16-SEP-1999;	99US-0154039.					
PR 20-SEP-1999;	99US-0154779.					
PR 22-SEP-1999;	99US-0155139.					
PR 23-SEP-1999;	99US-0155486.					
PR 24-SEP-1999;	99US-0155659.					
PR 28-SEP-1999;	99US-0156458.					
PR 29-SEP-1999;	99US-0156596.					
PR 04-OCT-1999;	99US-0157117.					
PR 05-OCT-1999;	99US-0157753.					
PR 06-OCT-1999;	99US-0157865.					
PR 07-OCT-1999;	99US-0158029.					
PR 08-OCT-1999;	99US-0158232.					
PR 12-OCT-1999;	99US-0158369.					
PR 13-OCT-1999;	99US-0159293.					
PR 13-OCT-1999;	99US-0159294.					
PR 13-OCT-1999;	99US-0159295.					
PR 14-OCT-1999;	99US-0159329.					
PR 14-OCT-1999;	99US-0159330.					
PR 14-OCT-1999;	99US-0159331.					
PR 14-OCT-1999;	99US-0159337.					
PR 14-OCT-1999;	99US-0159638.					
PR 18-OCT-1999;	99US-0159584.					
PR 21-OCT-1999;	99US-0160741.					
PR 21-OCT-1999;	99US-0160767.					
PR 21-OCT-1999;	99US-0160768.					
PR 21-OCT-1999;	99US-0160770.					
PR 21-OCT-1999;	99US-0160814.					
PR 21-OCT-1999;	99US-0160815.					
PR 22-OCT-1999;	99US-0160980.					
PR 22-OCT-1999;	99US-0160981.					
PR 22-OCT-1999;	99US-0160989.					
PR 25-OCT-1999;	99US-0161404.					
PR 25-OCT-1999;	99US-0161405.					
PR 25-OCT-1999;	99US-0161406.					
PR 26-OCT-1999;	99US-0161359.					
PR 26-OCT-1999;	99US-0161360.					
PR 26-OCT-1999;	99US-0161361.					
PR 28-OCT-1999;	99US-0161320.					
PR 28-OCT-1999;	99US-0161992.					
PR 28-OCT-1999;	99US-0161993.					
PR 29-OCT-1999;	99US-0162142.					
PR 04-AUG-1999;	99US-0147302.	QY	7	YKVADIGLAAGRKALDAENEMPGMLMRMRERYASKPLKRGARIAGCLMTVETAVLIET	66	
PR 05-AUG-1999;	99US-0147192.	DB	14	YKVKDMSQADFGRLLELAEEVEMPGLMACRTEFGSPQPFKGARITGSLHMTITQTAVLIET	73	
PR 05-AUG-1999;	99US-0147260.	QY	67	LVTILGAEVQWSSCNIFSTQNHAAAIKAGIIVYAWKGETDEEYLWCIEOTL-YFKDGPL	125	
PR 06-AUG-1999;	99US-0147303.	DB	74	LTALGAEVRCWSCNIFSTQDHAAAIARDAAAFPAWKGETLQEYWMCTERALDWDGPGGP	133	
PR 06-AUG-1999;	99US-0147416.	QY	126	NMILDDGGDLTNLIH-----TKYP	144	
PR 09-AUG-1999;	99US-0147493.	DB	134	DLIVDDGGDATLLIHGVKAEIIFKTCQVQPDPTSTDNPEFQIVLSIIKEGLQVDPKKYH	193	
PR 09-AUG-1999;	99US-0147935.	QY	145	QLLPGIRGISETTTGVHNLKYMANGILKVPAINVDSVTKSKFDNLYGCRSLIDGIK	204	
PR 10-AUG-1999;	99US-0148171.	DB	194	KMKERLVGVSEETTTGVKRLYQMQQNGTLLFPAINVDSVTKSKFDNLYGCRHSLPDGLM	253	
PR 11-AUG-1999;	99US-0148319.	QY	205	RATDVMIAGKAVAVAGYGVGKCAQALRGFCARVITIEDIPINALQAAAMEGYEVTWDE	264	
PR 12-AUG-1999;	99US-0148341.	DB	254	RATDVMIAGKAVICGYGVGKCAAAKTAGARVITIEDIPICALQALMEGLQVLTLED	313	
PR 13-AUG-1999;	99US-0148565.	QY	265	ACQEGNIFVTTTGCIDILGRHFEQKDDAIVCNIGHFDPDEIDVKWLN-NAVEKVNKP	323	
PR 13-AUG-1999;	99US-0149229.	DB	314	VVSEADIFVTTTGNKDIIMVDHMRKKNNAIVCNIGHFDEIDMLGLETPYGVKRIITKP	373	
PR 16-AUG-1999;	99US-0149368.	QY	324	QVDRY---RLKNRRRIILAEGLVNLGCAMGHPFVMSNSTFNOVMAQIELMTH--PDK	378	
PR 16-AUG-1999;	99US-0149368.	DB	374	QTRWVFPETKAG--IIVLAEGRLNLGCATGHPFVMSNSTFNOVIAQLELMNEKASCK	431	
PR 17-AUG-1999;	99US-0149175.	QY	379	YPGVGHFLPKLDEAVAENHGLKLVKLTEKQAQYLGMSCDGPFKPDHRY	432	
PR 18-AUG-1999;	99US-0149426.	DB	432	YEKKVYVLPKHLDEKVALHLGLGARLTKLSDQSDYVSIPIEGYPKPPHY	485	
PR 20-AUG-1999;	99US-0149722.	RESULT 10				
PR 20-AUG-1999;	99US-0149723.	AAG40085				
PR 23-AUG-1999;	99US-0149902.	ID AAG40085 standard; Protein; 508 AA.				
PR 23-AUG-1999;	99US-0149930.	XX				
PR 25-AUG-1999;	99US-0150566.	AC AAG40085;				
PR 26-AUG-1999;	99US-0150884.	XX				
PR 27-AUG-1999;	99US-0151065.	DT 18-OCT-2000 (first entry)				
PR 27-AUG-1999;	99US-0151066.	XX				
PR 30-AUG-1999;	99US-0151303.	DE Arabidopsis thaliana protein fragment SEQ ID NO: 49690.				
PR 31-AUG-1999;	99US-0151438.	XX				
PR 01-SEP-1999;	99US-0151930.	KW Protein identification; signal transduction pathway; metabolic pathway;				
PR 07-SEP-1999;	99US-0152363.	KW hybridisation assay; genetic mapping; gene expression control; promoter;				
PR 10-SEP-1999;	99US-0153070.	XX				
PR 13-SEP-1999;	99US-0153758.	XX				
PR 15-SEP-1999;	99US-0154018.	XX				
PR 16-SEP-1999;	99US-0154039.	XX				
PR 20-SEP-1999;	99US-0154779.	XX				
PR 22-SEP-1999;	99US-0155139.	OS Arabidopsis thaliana.				
PR 23-SEP-1999;	99US-0155486.	XX				
PR 24-SEP-1999;	99US-0155659.	PN EP1033405-A2.				
PR 28-SEP-1999;	99US-0156458.	PD 06-SEP-2000.				
PR 29-SEP-1999;	99US-0156596.	XX				
PR 04-OCT-1999;	99US-0157117.	XX				
PR 05-OCT-1999;	99US-0157753.	XX				
PR 06-OCT-1999;	99US-0157865.	XX				
PR 07-OCT-1999;	99US-0158029.	XX				
PR 08-OCT-1999;	99US-0158232.	XX				
PR 12-OCT-1999;	99US-0158369.	XX				
PR 13-OCT-1999;	99US-0159293.	XX				
PR 13-OCT-1999;	99US-0159294.	XX				
PR 13-OCT-1999;	99US-0159295.	XX				
PR 14-OCT-1999;	99US-0159329.	XX				
PR 14-OCT-1999;	99US-0159330.	XX				
PR 14-OCT-1999;	99US-0159331.	XX				
PR 14-OCT-1999;	99US-0159337.	XX				
PR 14-OCT-1999;	99US-0159638.	XX				
PR 18-OCT-1999;	99US-0159584.	XX				
PR 21-OCT-1999;	99US-0160741.	XX				
PR 21-OCT-1999;	99US-0160767.	XX				
PR 21-OCT-1999;	99US-0160768.	XX				
PR 21-OCT-1999;	99US-0160770.	XX				
PR 21-OCT-1999;	99US-0160814.	XX				
PR 21-OCT-1999;	99US-0160815.	XX				
PR 22-OCT-1999;	99US-0160980.	XX				
PR 22-OCT-1999;	99US-0160981.	XX				
PR 22-OCT-1999;	99US-0160989.	XX				
PR 25-OCT-1999;	99US-0161404.	XX				
PR 25-OCT-1999;	99US-0161405.	XX				
PR 25-OCT-1999;	99US-0161406.	XX				
PR 26-OCT-1999;	99US-0161359.	XX				
PR 26-OCT-1999;	99US-0161360.	XX				
PR 26-OCT-1999;	99US-0161361.	XX				
PR 28-OCT-1999;	99US-0161320.	XX				
PR 28-OCT-1999;	99US-0161992.	XX				
PR 28-OCT-1999;	99US-0161993.	XX				
PR 29-OCT-1999;	99US-0162142.	XX				

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0130891.
PR 21-JUL-1999; 99US-0131449.
PR 21-JUL-1999; 99US-0132048.
PR 21-JUL-1999; 99US-0132407.
PR 22-JUL-1999; 99US-0132484.
PR 22-JUL-1999; 99US-0132485.
PR 22-JUL-1999; 99US-0132486.
PR 22-JUL-1999; 99US-0132487.
PR 23-JUL-1999; 99US-0132863.
PR 23-JUL-1999; 99US-0134256.
PR 23-JUL-1999; 99US-0134218.
PR 23-JUL-1999; 99US-0134219.
PR 23-JUL-1999; 99US-0134221.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

```

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 56.9%; Score 1292; DB 21; Length 508;
Best Local Similarity 56.3%; Pred. No. 8.3e-125;
Matches 267; Conservative 55; Mismatches 102; Indels 50; Gaps 6;

QY 7 YKVADIGLAAGRKALDIAENEMPGIMRMREYSASKPLKGARIAGCLHMTVETAVLIET 66
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
37 YKVQMSQADFGRLELEAEVEMPGIMACRTEFGSPQPFKARITGSLHMTIQTAVLIET 96
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 67 LVTLAGAEVOWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEBYLWCIBQTL-YFKDGPL 125
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
97 LTAGLAEVRWCSNIFSTQDHAAAAIARDASAFAWKGTELQBYWWTCTERALDWMGPGGP 156
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 126 NMILDGGDLTNLIH-----TKYP 144
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
157 DLIVDDGGDTLLIHGKVAEEIFEKTQGVDPDPTDNPFEFQIVLSIIKEGLQVDPKKYH 216
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 145 QLLPGIRGISEETTTGVHNLKMWANGILKVPAINVNDSTKSKPDNLGCGRESLIDGK 204
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
217 KMKERLVGVSEETTTGVKRLYQMQNGTLFPAINVNDSTKSKPDNLGCGRHSPLDGLM 276
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 205 RATDVMIAKVAVAGYGVGKCAQALRGFGARVITTEIDIPINALQAAMEGEVETVMD 264
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
277 RATDVMIAKVAVICGYGVGKCAAAKMTAGARVITTEIDIPICALQALMEGLQVLTLED 336
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 265 AQOEGNIFVTTTCIDIIILGRHFEQMKDDAIVCNIGHFDEIDVKWLN-NAVEKNIKP 323
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
337 VVSEADIFVTTTGNKDIIMVDMRMKNNNAIVCNIGHFDEIDMLGLETYPGVKRTIKP 396
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 324 QVDRY---RLKNGRRILLAEGLVNLGCAMGHPSFVMSNSFTNQVMAQIELWTH--PDK 378
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
397 QTRDWRVFPETKAG--LIVLAEGRLMNLGCATGHPSFVMSCSFTNQVIAQLELWNEKASGK 454
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 379 YPVGVHFLPKLDEAAVAEHLGKLNKLTKEQQAQYLGMSCDGPFKPDHYRY 432
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
455 YEKVVVLPKHLDEKVALHLLGKLGARLTKLSKQSDYVSIPIEGYKPPHYRY 508
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 11
AAW01458
ID AAW01458 standard; Protein; 485 AA.
XX
AC AAW01458;
XX
DT 22-FEB-1997 (first entry)
DE
DE Asparagus S-adenosyl-L-homocysteine hydrolase.
XX
KW S-adenosyl-L-homocysteine hydrolase; SHH; promoter; monooct; dicot;
XX transgenic plant; disease resistance; pathogen resistance.
OS Asparagus officinalis.

```

```

XX Key Location/Qualifiers
FH Region 150..190
FT /note= "region found in SHH of photosynthetic
FT species"
FT Binding-site 263..294
FT /label= NAD+_binding_site
XX
PN WO9632488-A1.
XX
PD 17-OCT-1996.
XX
PF 10-APR-1996; 96WO-GB00882.
XX
PR 10-APR-1995; 95GB-0007381.
XX
PA (ZONE ) ZENECA LTD.
XX
PI Draper J, Greenland AJ, Skipsey M, Warner S;
XX WPI; 1996-477138/47.
DR N-PSDB; AAT44513.
XX
PT S-adenosyl-L-homocysteine hydrolase promoter - used for driving
PT expression of effector genes, such as pathogen resistance genes, in
PT transgenic plants
XX
XX Example 1; Fig 1; 57pp; English.
XX
CC Asparagus S-adenosyl-L-homocysteine hydrolase (SHH) (AAW01458)
CC catalyses the reversible hydrolysis of S-adenosyl-L-homocysteine
CC to adenosine and homocysteine. It contains an extra stretch of
CC amino acid residues (positions 150-190) previously found in other
CC photosynthetic species, parsley and Rhodospirillum rubrum, but not
CC in SHHs from non-photosynthetic species. A cDNA sequence (AAT44513)
CC coding for the asparagus was used to identify the Arabidopsis
CC thaliana SHH gene (AAT44515) and promoter (AAT44514), useful for
CC expression of effector genes in transgenic plants.
XX
XX Sequence 485 AA;
XX
Query Match 56.8%; Score 1289; DB 17; Length 485;
Best Local Similarity 56.1%; Pred. No. 1.6e-124;
Matches 265; Conservative 55; Mismatches 106; Indels 46; Gaps 5;

QY 7 YKVADIGLAAGRKALDIAENEMPGIMRMREYSASKPLKGARIAGCLHMTVETAVLIET 66
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
14 YKVQMSQADFGRLELEAEVEMPGIMACRTEFGSPQPFKARITGSLHMTIQTAVLIET 73
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 67 LVTLAGAEVOWSSCNIFSTQNHAAAAIAKAGIPVYAWKGETDEBYLWCIBQTL-YFKDGPL 125
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
74 LTAGLAEVRWCSNIFSTQDHAAAAIARDASAFAWKGTELQBYWWTCTERALDWMGPGGP 133
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 126 NMILDGGDLTNLIH-----TKYP 144
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
134 DLIVDDGGDTLLIHGKVAEEIFEKTQGVDPDPTDNPFEFQIVLSIIKEGLQVDPKKYR 193
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 145 QLLPGIRGISEETTTGVHNLKMWANGILKVPAINVNDSTKSKPDNLGCGRESLIDGK 204
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
194 KMKDRIVGVSEETTTGVKRLYQMQNGTLFPAINVNDSTKSKPDNLGCGRHSPLDGLM 253
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 205 RATDVMIAKVAVAGYGVGKCAQALRGFGARVITTEIDIPINALQAAMEGEVETVMD 264
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
254 RATDVMIAKVAVCGYGVGKCAAAKMTAGARVITTEIDIPICALQALMEGLQVLTLED 313
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 265 AQOEGNIFVTTTCIDIIILGRHFEQMKDDAIVCNIGHFDEIDVKWLN-NAVEKNIKP 323
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
314 VVSEADIFVTTTGNKDIIMVDMRMKNNNAIVCNIGHFDEIDMLGLETYPGVKRTIKP 373
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY 324 QVDRY---RLKNGRRILLAEGLVNLGCAMGHPSFVMSNSFTNQVMAQIELWTH--PDKYP 380
DB ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
374 QTRDWRVFPETNTGIIIVLAEGRLMNLGCATGHPSFVMSCSFTNQVIAQLELWNEKASGYE 433
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

```

QY 381 VGVHFLPKLDEAVAHGLKLVKLTKEQAQYLGMSCDGPFKPDHYRY 432
 Db : ||| ||| ||| ||| : ||| : ||| |||
 434 KKYVLPKHLDERVAALHLGLGAKLTKLSPSQADVISVPIEGPKPPHYRY 485

RESULT 12

AAG41407
 ID AAG41407 standard; Protein; 485 AA.

XX AC AAG41407;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 51514.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 23-APR-1999; 99US-0130891.

XX PR 28-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 06-MAY-1999; 99US-0132487.

XX PR 11-MAY-1999; 99US-0132863.

XX PR 14-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 01-JUL-1999; 99US-0141287.

PR 30-JUN-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 28-JUL-1999; 99US-0145919.

PR 02-AUG-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	23-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0152370.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161320.
PR	28-OCT-1999;	99US-0161392.
PR	28-OCT-1999;	99US-0161393.
PR	29-OCT-1999;	99US-0162142.
PR	29-OCT-1999;	99US-0162143.
Query Match 56.1%; Score 1274; DB 21; Length 485;		
Best Local Similarity 55.8%; Pred. No. 5.7e-123;		
Matches 262; Conservative 57; Mismatches 107; Indels 46; Gaps 5;		
QY	7	YKVADTGLAARWKAIDIAENEMPGLRMRERYASAKPLKGARIAGLHMTVETAVLIET 66
Db	14	YKVQNSQADFGRLLELEAVEVPGVLCVTEFGPSQPLKGRITGSLHMTQTAVLIET 73
QY	67	LVTLGEVQWSSCNIFSTQNHAAAAIAKAGIPYAWKGETDEBYLWCIBQTL-YFKDGPL 125
Db	74	LTAALGAEVWCSNIFSTQDAAAAIARDSAAVFAWKGETLQBYWVWCTERALDWGPGGGP 133

QY	126	NMILDDGGDLTNLIH-----	-----TKYP 144
Db	134	DLIVDDGGDATLLIHGVKAEEIFAKNGTFFDPPTSTDNPEFOIVLSIIKDGLOVDPKYYH 193	
QY	145	QLLPGIRGISEETTGVHNLVYKWMANGILKVPAINVNDVSVTKSKFDNLYGCCRESLIDGIK 204	
Db	194	KMKERLVGVSEETTGVKELYQMETGALLFPAINVNDVSVTKSKFDNLYGCRHSLPDGLM 253	
QY	205	RATDVMIAGKAVAVAGYGDVHGKCAQALRGFGARVIITEIDPINALQAAAMEGVYVTTMOE 264	
Db	254	RATDVMIAGKAVICGYGDVHGKCAAMKTAGARVITVEIDPICALQALMEGLQVLTLED 313	
QY	265	ACQEGNIFVTTTGCIDIILGRHFEQMKDDAIVCNIGHDFDEIDVKWLN-NAVEKVNIKP 323	
Db	314	VVSEADIFCTTTGNKDIIMVDHMRKNNNAIVCNIGHDFDEIDMLGLETYPGVKRITIKP 373	
QY	324	QVDRYRLKN-GRRIILLAEGRVLNLCAMGHPHSFVMSNSFTNQVMAQIELWTHPD--KYP 380	
Db	374	QTRWVFPDTSGLIIVLAEGRLMNLGCATGHPHSFVMSCSFTNQVIAQLELWNEKSSGKYE 433	
QY	381	GVHFLPKKLDVAEVAHLKLVNLTXTLTKQAQYVLGMSCDGPPKPDHYRY 432	
Db	434	KKVYVLPKHLDEKVAALHLKLGARLTKLTQSDVYSIPVEGPKPVHYRY 485	
RESULT 13			
AAG41406			
ID	AAG41406 standard; Protein; 497 AA.		
XX	AAG41406;		
XX	DT 18-OCT-2000 (first entry)		
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 51513.		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 51513.		
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX	termination sequence.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-0301439.		
XX	99US-0121825.		
PR	99US-0123180.		
PR	99US-0123548.		
PR	99US-0125788.		
PR	99US-0126264.		
PR	99US-0126785.		
PR	99US-0127462.		
PR	99US-0128234.		
PR	99US-0128714.		
PR	99US-0129845.		
PR	99US-0130077.		
PR	99US-0130449.		
PR	99US-0130510.		
PR	99US-0130891.		
PR	99US-0131449.		
PR	99US-0132048.		
PR	99US-0132407.		
PR	99US-0132484.		
PR	99US-0132485.		
PR	99US-0132486.		
PR	99US-0132487.		
PR	99US-0132863.		
PR	99US-0134256.		
PR	99US-0134218.		
PR	99US-0134219.		

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

[illegible]

PR	16-JUL-1999;	99US-0144086;
PR	19-JUL-1999;	99US-0144325;
PR	19-JUL-1999;	99US-0144331;
PR	19-JUL-1999;	99US-0144332;
PR	19-JUL-1999;	99US-0144333;
PR	19-JUL-1999;	99US-0144334;
PR	19-JUL-1999;	99US-0144335;
PR	20-JUL-1999;	99US-0144352;
PR	20-JUL-1999;	99US-0144632;
PR	20-JUL-1999;	99US-0144884;
PR	21-JUL-1999;	99US-0144814;
PR	21-JUL-1999;	99US-0145086;
PR	21-JUL-1999;	99US-0145088;
PR	22-JUL-1999;	99US-0145085;
PR	22-JUL-1999;	99US-0145087;
PR	22-JUL-1999;	99US-0145089;
PR	22-JUL-1999;	99US-0145192;
PR	23-JUL-1999;	99US-0145145;
PR	23-JUL-1999;	99US-0145218;
PR	23-JUL-1999;	99US-0145224;
PR	26-JUL-1999;	99US-0145276;
PR	27-JUL-1999;	99US-0145913;
PR	27-JUL-1999;	99US-0145918;
PR	27-JUL-1999;	99US-0145919;
PR	28-JUL-1999;	99US-0145951;
PR	02-AUG-1999;	99US-0146386;
PR	02-AUG-1999;	99US-0146388;
PR	02-AUG-1999;	99US-0146389;
PR	03-AUG-1999;	99US-0147038;
PR	03-AUG-1999;	99US-0147204;
PR	04-AUG-1999;	99US-0147302;
PR	05-AUG-1999;	99US-0147192;
PR	05-AUG-1999;	99US-0147260;
PR	06-AUG-1999;	99US-0147303;
PR	06-AUG-1999;	99US-0147416;
PR	09-AUG-1999;	99US-0147493;
PR	09-AUG-1999;	99US-0147935;
PR	10-AUG-1999;	99US-0148171;
PR	11-AUG-1999;	99US-0148319;
PR	12-AUG-1999;	99US-0148341;
PR	13-AUG-1999;	99US-0148565;
PR	13-AUG-1999;	99US-0148684;
PR	16-AUG-1999;	99US-0148686;
PR	17-AUG-1999;	99US-0149175;
PR	18-AUG-1999;	99US-0149426;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149723;
PR	22-AUG-1999;	99US-0149929;
PR	23-AUG-1999;	99US-0149902;
PR	27-AUG-1999;	99US-0149930;
PR	28-AUG-1999;	99US-0150566;
PR	26-AUG-1999;	99US-0150884;
PR	27-AUG-1999;	99US-0151065;
PR	27-AUG-1999;	99US-0151066;
PR	27-AUG-1999;	99US-0151080;
PR	30-AUG-1999;	99US-0151303;
PR	31-SEP-1999;	99US-0151438;
PR	01-SEP-1999;	99US-0151930;
PR	07-SEP-1999;	99US-0152363;
PR	10-SEP-1999;	99US-0153070;
PR	13-SEP-1999;	99US-0153758;
PR	15-SEP-1999;	99US-0154018;
PR	16-SEP-1999;	99US-0154039;
PR	20-SEP-1999;	99US-0154779;
PR	22-SEP-1999;	99US-0155139;
PR	23-SEP-1999;	99US-0155486;
PR	24-SEP-1999;	99US-0155659;
PR	28-SEP-1999;	99US-0156458;
PR	29-SEP-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	05-OCT-1999;	99US-0157753;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158029;

PR	08-OCT-1999;	99US-01582312
PR	12-OCT-1999;	99US-01583369
PR	13-OCT-1999;	99US-01582923
PR	13-OCT-1999;	99US-0159294
PR	13-OCT-1999;	99US-0159295
PR	13-OCT-1999;	99US-0159295
PR	14-OCT-1999;	99US-0159295
PR	14-OCT-1999;	99US-01593330
PR	14-OCT-1999;	99US-01593330
PR	14-OCT-1999;	99US-0159637
PR	14-OCT-1999;	99US-0159638
PR	18-OCT-1999;	99US-0159584
PR	21-OCT-1999;	99US-0160741
PR	21-OCT-1999;	99US-0160767
PR	21-OCT-1999;	99US-0160768
PR	21-OCT-1999;	99US-0160770
PR	21-OCT-1999;	99US-0160814
PR	21-OCT-1999;	99US-0160815
PR	22-OCT-1999;	99US-0160980
PR	22-OCT-1999;	99US-0160981
PR	22-OCT-1999;	99US-0160989
PR	25-OCT-1999;	99US-0161404
PR	25-OCT-1999;	99US-0161405
PR	25-OCT-1999;	99US-0161406
PR	26-OCT-1999;	99US-0161359
PR	26-OCT-1999;	99US-0161360
PR	26-OCT-1999;	99US-0161361
PR	28-OCT-1999;	99US-0161320
PR	28-OCT-1999;	99US-0161992
PR	28-OCT-1999;	99US-0161993
PR	29-OCT-1999;	99US-0162144

Query Match 56.0%; Score 1272; DB 21; Length 467;

Best Local Similarity 56.4%; Pred. No. 8.7e-123;

Matches 263; Conservative 54; Mismatches 99; Indels 50; Gaps 6;

[illegible]

RESULT 15
AAG41408
ID AAG4
XX
AC AAG4

XX	18-OCT-2000 (first entry)	99US-0139750
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 51515.	99US-0139763
DE	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.	99US-0139817
XX	Arabidopsis thaliana.	99US-0139899
XX	EP1033405-A2.	99US-0140353
PD	06-SEP-2000.	99US-0140354
XX	25-FEB-2000; 2000EP-0301439.	99US-0140695
XX	25-FEB-1999; 99US-0121825.	99US-0140991
PR	05-MAR-1999; 99US-0123180.	99US-0141287
PR	09-MAR-1999; 99US-0123548.	99US-0141842
PR	23-MAR-1999; 99US-0125788.	99US-0142055
PR	25-MAR-1999; 99US-0126264.	99US-0142390
PR	29-MAR-1999; 99US-0126785.	99US-0142803
PR	01-APR-1999; 99US-0127462.	99US-0142920
PR	06-APR-1999; 99US-0128234.	99US-0142977
PR	08-APR-1999; 99US-0128714.	99US-0143542
PR	16-APR-1999; 99US-0129845.	99US-0143624
PR	19-APR-1999; 99US-0130077.	99US-0143909
PR	21-APR-1999; 99US-0130449.	99US-0144297
PR	23-APR-1999; 99US-0130510.	99US-0144331
PR	28-APR-1999; 99US-0130891.	99US-0144332
PR	30-APR-1999; 99US-0131449.	99US-0144333
PR	30-APR-1999; 99US-0132048.	99US-0144334
PR	04-MAY-1999; 99US-0132407.	99US-0144335
PR	05-MAY-1999; 99US-0132484.	99US-0144352
PR	06-MAY-1999; 99US-0132485.	99US-0144632
PR	07-MAY-1999; 99US-0132486.	99US-0144884
PR	11-MAY-1999; 99US-0132863.	99US-0144814
PR	14-MAY-1999; 99US-0134256.	99US-0145086
PR	14-MAY-1999; 99US-0134218.	99US-0145088
PR	14-MAY-1999; 99US-0134219.	99US-0145085
PR	14-MAY-1999; 99US-0134221.	99US-0145087
PR	18-MAY-1999; 99US-0134370.	99US-0145089
PR	19-MAY-1999; 99US-0134768.	99US-0145192
PR	20-MAY-1999; 99US-0134941.	99US-0145198
PR	21-MAY-1999; 99US-0135124.	99US-0145919
PR	24-MAY-1999; 99US-0135353.	99US-0145951
PR	25-MAY-1999; 99US-0135629.	99US-0146386
PR	27-MAY-1999; 99US-0136021.	99US-0146388
PR	28-MAY-1999; 99US-0136392.	99US-0147038
PR	01-JUN-1999; 99US-0136782.	99US-0147204
PR	03-JUN-1999; 99US-0137222.	99US-0147302
PR	04-JUN-1999; 99US-0137528.	99US-0147192
PR	07-JUN-1999; 99US-0137502.	99US-0147260
PR	08-JUN-1999; 99US-0137724.	99US-0147303
PR	10-JUN-1999; 99US-0138094.	99US-0147416
PR	10-JUN-1999; 99US-0138540.	99US-0147493
PR	14-JUN-1999; 99US-0138847.	99US-0147935
PR	16-JUN-1999; 99US-0139119.	99US-0148171
PR	17-JUN-1999; 99US-0139452.	99US-0148319
PR	17-JUN-1999; 99US-0139453.	99US-0148341
PR	18-JUN-1999; 99US-0139457.	99US-0148565
PR	18-JUN-1999; 99US-0139458.	99US-0148684
PR	18-JUN-1999; 99US-0139459.	99US-0149368
PR	18-JUN-1999; 99US-0139460.	99US-0149175
PR	18-JUN-1999; 99US-0139461.	99US-0149266
PR	18-JUN-1999; 99US-0139462.	99US-0149426
PR	18-JUN-1999; 99US-0139463.	99US-0149722
PR	18-JUN-1999; 99US-0139463.	99US-0149723
PR	18-JUN-1999; 99US-0139463.	99US-0149729
PR	18-JUN-1999; 99US-0139463.	99US-0149902
PR	18-JUN-1999; 99US-0139463.	99US-0149930
PR	18-JUN-1999; 99US-0139463.	99US-0150566
PR	18-JUN-1999; 99US-0139463.	99US-0150884

```
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 55.2%; Score 1254; DB 21; Length 467;
Best Local Similarity 55.6%; Pred. No. 6.4e-121;
Matches 258; Conservative 56; Mismatches 104; Indels 46; Gaps 5;
```

```
Qy 15 AAMGRKALDIENEMPCGLMWRERYASKPLKGRIAGCLHMTVETAVLIETLVTGAEV 74
Db 4 ADFGRLEIAEVEMLPGLVSCVTEFGPSQPLKGARITGSLHMTIQTAVLIETLTALGAEV 63

Qy 75 QWSSCNIFSTQNHAAAIKAGIPVYAKGSTDDEYLWCIEQTL-YFKDGPLNMLDDGG 133
Db 64 RWCSCNIFSTQDHAAAIARDAARVAFKAGETLQEYVWCTERALDWGPGGPDLLIVDDGG 123

Qy 134 DLTNLIH-----TKYPQLPGIRG 152
Db 124 DATLLIHEGVKAEIIFAKNGTFDPDTSTDNPEFOIVLSIIKQGLQVDPKKYHKMKERLVG 183

Qy 153 ISEETTTGVNLYKMWANGILKVPAINVNDSVTKSKFDNLYGCRSLIDGIKRATDVNIA 212
Db 184 VSEETTTGVKRLYQMQETGALLFPAINVNDSVTKSKFDNLYGCRHSLPDLGMRATDVNIA 243

Qy 213 GKAVVAGYGVGKGCQAALRGFGARVITTEIDPINALQAAMEGYEVTWDEACQEGNIF 272
```

Search completed: December 30, 2003, 10:40:35
Job time : 48 secs

3 PAGE BLANK (USPTO)